VOL. 47, NO. 1 JANUARY, 1940

THE PSYCHOLOGICAL REVIEW

THE TIME ORDER ERROR IN SUCCESSIVE JUDGMENTS AND IN REFLEXES. III. TIME ERROR THEORIES

BY HELEN PEAK

Randolph-Macon Woman's College

The results of an experiment which has been reported elsewhere (**25, 26**) are summarized briefly at this time in order that the various theories of the time order error may be evaluated in the light of these new findings. The loudness of paired tones was judged on a 9-step absolute scale, and lid reflexes to the same stimuli were recorded photographically. Members of a pair of tones were spaced at 177 ms. or 285 ms. intervals. Single tones were interspersed in the series. The intensity of the first of the paired tones and of the single tones varied; that of the second tone of the pair remained constant. Results from 5 subjects have shown: (1) that the size of the judged loudness of, and reflexes to, the second tone of a pair [1] decreases with increase in the intensity of the first stimulus, except at the largest intensity of S_1 where the J_2 function reverses itself; (2) that following the largest two intensities of S_1 at an interval of 285 ms., J_2 and R_2 show less decrement than at 177 ms. At the smaller intensities, however, inhibitory effects tend to increase at the longer interval; (3) that J_1 paired is greater than judgments of single stimuli, except at the largest intensity of S_1 where the results are not uniform; (4) that low positive correlations exist between J_1 and J_2, between R_1 and R_2, and between J_2 and R_2, when these responses follow a first stimulus of either of the two largest intensities. The correlations are near zero or are very slightly

[1] As before, the symbol S_1 represents the first stimulus of a pair, S_2 the second, R_1 and R_2, the reflexes to the first and second stimuli, and J_1 and J_2, judgments of the first and second stimuli.

negative when S_1 is weak; (5) that positive time errors occur at 177 ms. and at 285 ms. intervals and that this phenomenon depends both on decrease in the judged loudness of a stimulus when it is preceded by another stimulus and upon enhancement of the judged loudness of the first stimulus by pairing it with a second stimulus.

These results leave no doubt of the effect of each of the paired stimuli on the judgment of the other and it is the purpose of the present paper to consider the origin of these effects. Five possible sources of modification of judgments by pairing have been set down below and will be considered in detail as potential factors in any time error theory which attempts to account for changes in both J_1 and J_2: (1) Progressive changes with time in the excitation processes resulting from the first stimulus. The course of this change helps determine the level of the 'trace' of S_1 when S_2 is presented. (2) Processes set up by the first stimulus which may affect the reaction to the second stimulus. These effects may be facilitative or inhibitory, and conceivably they may appear without practice or instructions, or they may result from practice in the paired stimulus situation in which the first stimulus becomes established as a signal for the appearance of the second. (3) Reciprocal interaction between traces of the first and second stimuli, including (*a*) contrast effects or (*b*) tendencies toward levelling and equilibrium. (4) Possible modification of J_1 by after-effects of S_2. (5) Interactions of one or both traces with background processes in the organism.

I. Trace Theories

The assumption of a gradual fading of the trace left by the first stimulus of a pair constituted the chief feature of the early attempts to explain time order errors.[2] Though Köhler (**14**), Lauenstein (**16**), Borak (**2**), and others have shown this theory inadequate to account for all the data of

[2] The distinction between the fading memory image and a neurological trace of decreasing intensity need not concern us here. Experiments have shown that the image is not always present, but it is impossible to determine the neurological nature of the after-effects. See Needham (**23**) for a summary of the theories.

comparison judgments, it is clear, and practically all writers agree [3] that some after-effects of the stimulus must serve as a factor in the judgment of the absolute or relative intensity of the first stimulus. Such a process is doubtless of importance for the negative time error, but it has always seemed to leave little room for the existence of a positive error.

Köhler's notion (14) of a stimulus trace of S_1, which attains its maximum value slowly before beginning to decline, constitutes one attempt to circumvent this difficulty. He suggested that if the second stimulus was presented during the rising phase of the first, a step down in excitation value would result, a falling from high to low potential which would determine a judgment of 'less.' It is not clear how this theory can serve to account for the positive error unless it be assumed further that judgment of S_1 and S_2 occurs so soon after the presentation of S_2 that the resulting S_2 excitation process has not had time to recruit to the maximum value which it will develop in time. This follows from the fact that if both S_1 and S_2 have fully recruited when judgment occurs, the S_1 process should always be at a less intense level than S_2 because it will have had a longer time in which to fade. The theory with this addition would run thus: If S_1 and S_2 are objectively equal, if the S_1 trace builds up gradually and then begins to fade, if the S_2 process takes the same course as S_1, and if, finally, judgments take place immediately after S_2 is presented when S_2 excitation is just beginning to build up, then S_2 may be judged smaller than S_1 and a positive error should always result. It is not apparent why a negative error should ever occur. On the other hand, if it be assumed that the judgment takes place any time after recruitment of the S_2 process is complete, the occurrence of a positive error is difficult to explain, for under these circumstances both processes, having developed fully, will begin to fade and the first, having faded for a longer time, should be judged weaker. Thus neither assumption about time of judgment can explain both negative and positive errors.

[3] Wheeler and Perkins (32, p. 391) constitute a possible exception. See Koffka (13, p. 445) for discussion and criticism of this view.

Another objection to this theory may be mentioned. If it should be admitted that judgments of S_1 and S_2 depend on a comparison process occurring so soon after the presentation of S_2 that this latter process is not fully developed and is therefore under-estimated, it would still be necessary to account for the appearance of decrement in J_2 when S_1 is judged weaker than S_2, and for the increment in J_1 under these circumstances. Obviously Köhler's theory deals with a different experimental set-up and consequently makes no attempt to explain how judgments of both stimuli could behave as they do. It is difficult to see, however, in what manner this or the other fading trace theories could be adapted to explain the present facts.

II. Possible Effects of S_1 Processes on Reactions to S_2

a. Inhibition.—Kreezer (**15**) appears to be the only writer who has made a definite attempt to relate time order errors to reflex inhibition and he has been concerned chiefly with the possible neurological locus of such processes. Müller and Martin (**20**) and later Köhler (**14**) may have been near to an inhibition theory in the attempted explanation of time errors by what they called 'fatigue.' Müller suggested fatigue as a factor in the negative error for lifted weights; Köhler proposed it as one alternative to explain the positive error with sound stimuli. Lehmann (**17**), however, criticized Müller's use of the concept as not applicable to auditory data, and Köhler himself dismissed the fatigue concept as unsatisfactory because of the supposed absence of the positive error at intervals shorter than $1\frac{1}{2}$ sec.

In the first paper of this series (**25**) certain generalizations have been made regarding inhibition of responses to the second stimulus of a pair in spinal reflexes, in the lid reflex, and in judgments of paired tones, and it has been suggested that the factors responsible for the appearance of inhibition in J_2 must account in part for the positive time error observed. It is unnecessary at this time to repeat the justification offered for the particular generalizations made, but the fact that the time order error in judgments has been shown to

bear certain similarities to the behavior of reflexes suggests that similar intervening variables may be at work in producing variation in these functions. Whether the processes are really the same and what they are cannot now be determined. The data offered and the comparisons made, simply pose the important question whether in the light of similarity of observed functions, similar neural processes may not be at work at these different levels of behavior. When the neural events which are responsible for inhibitory effects in any of these situations are eventually understood, they should be considered as factors of possible importance in explaining the others, but only experimental test will show whether they are actually determining factors in each case.

In the first paper it has been shown that the following experimental observations may be subsumed under three generalizations regarding the relation of stimulus intensity to the course of the development of inhibitory processes in spinal reflexes, lid reflexes, and judgments: (1) the decrement of J_2 and R_2 following S_1, (2) the increase in decrement with increase in intensity of S_1, (3) the effects of stimulus interval on J_2 and R_2 at different intensities of S_1. It is possible that these generalizations must be supplemented in order to account for the terminal rise in the J_2 function. It is certain that they must be supplemented in order to explain: (1) the enhancement of J_1 in the paired stimulus situation as compared to judgments of the same stimuli alone, (2) the positive correlation between J_1 and J_2 and the increase in this value with increasing intensity of S_1. In other words, principles other than those governing inhibition of J_2 must be added in order to account completely for the positive time error, inasmuch as changes in J_1 are involved in that error. Possible explanations of these changes are considered below.

b. Facilitation.—It has been demonstrated repeatedly that inhibitory and facilitatory processes are intimately related, so that if inhibition analogous to that appearing in other reactions is characteristic of judgment processes, facilitation would also be expected to occur. In fact, as we have pointed out in the first paper, such a process may be needed in order

to account for the terminal rise in the J_2 curve, should it be discovered that there are no grounds for the assumption that inhibition set up by a very intense S_1 has fallen off by the time S_2 appears.[4] It is also interesting to speculate whether such facilitation of J_2 by processes arising from S_1 may be involved in the appearance of the negative time error at some of the longer intervals, or in such a result as Wada's (**31**), who finds cycles of positive and negative time errors, with a positive error appearing at intervals as long as one minute. It is tempting to look for similarities in these functions and in some of the scattered and, as yet rare, physiological studies of cycles of facilitation and inhibition at longer intervals. Dusser de Barenne and McCulloch (**5**), for example, report fluctuations in excitability of isolated points on the motor cortex when faradic stimuli are employed. The complete range of time intervals used is not clear, but a period of decrement definitely followed one of facilitation. The point of maximal 'extinction' (*i.e.*, decrement) varied with stimulus characteristics and with cortical conditions, the optimal interval becoming longer with increased narcosis. Under 'moderate' narcosis, maximum 'extinction' appeared at a 13 sec. interval. Fulton (**8**, p. 408) reports an optimal interval of 4 sec. under a local anæsthetic. In a normal waking state this interval might well be shortened to the value characteristic of positive time error results with facilitation cycles changing in a corresponding fashion. These facts should certainly be kept in mind by the psychologist interested in the time order error.

c. Modification of J_2 by after effects of S_1, developed through practice.—It is impossible to postulate precisely the extent to which this sort of facilitation and inhibition would be

[4] It is likely that delay path activity is involved in the facilitation of one process by another and that a succession of neural impulses set up by a stimulus impinges on a given center by way of paths of various degrees of circuitousness, of different lengths and times of recovery, producing thereby cycles of facilitation and inhibition. Lorento de Nó has shown (**18**) the importance of internuncial pathways for both inhibition and facilitation. His work, that of Sherrington and his collaborators (**3**), and of Dusser de Barenne (**5**) and others must be considered in any attempt at a neurological theory of the time order error. This will not be attempted at this time.

different from that in (*a*) and (*b*), for the effect in both cases is assumed to be set up by S_1. It is possible, however, to make a distinction on operational grounds between effects of S_1 which occur previous to practice, and those which appear after practice or as a result of instructions in a situation where one stimulus follows consistently upon another. As pointed out in the first paper, this situation always affords the possibility that S_1 will become a signal for the appearance of S_2, so that the second stimulus will be anticipated and the judgment modified (either facilitated or inhibited). Culler (**4**) has put forward an idea which appears to be related to this when he attempts to explain the preponderance of correct judgments in cases where the variable follows the standard stimulus. He states that if two stimuli have been frequently given together, they become 'assimilated' so that when one is offered, it tends to arouse an impression of the other stimulus. Therefore, if the standard precedes, there is a tendency for all the variable patterns to be activated equally and not to any appreciable degree. When the variable itself is applied, a disparity results and discrimination is easy. Though Culler considers this a Gestalt doctrine, the phenomenon evidently depends upon building up an anticipatory reaction by practice which may also be said to resemble conditioning.

The so-called 'set' theories of the time error will be included here since they assume that a preparatory state is established in the organism by the first stimulus of a pair. It is doubtful whether the protagonists of the theories to be discussed have always been concerned with the factor of practice, but there is no doubt that they have emphasized the role of an after-effect of the first stimulus, which in fact, must have been influenced by the consistency of the paired stimulus relationship. Müller and Schumann (**21**) proposed that as a result of lifting the first weight of a pair, the subject adjusted the force of his lifting movement to the weight and used this first lifting as a basis of preparation for the second. If the second came up easily, it was judged lighter. Lehmann's (**17**) theory of the positive time error was similar except for his statement that the first weight was lifted more

rapidly and with more kinæsthesis than necessary, while the second lifting was adjusted more perfectly to the required force (when the weights were equal) with the consequence that the second was judged lighter. Borak (**2**) also proposed a set theory, assuming a residue of muscle tonus from lifting a first weight which was said to combine with and enhance the impression of the second weight. These particular sets, relating as they do to the weight lifting situation, do not appear to lend themselves to explanations of the judgments of paired sounds, but there may well be other kinds of sets established by S_1, as suggested above. This possibility led us to expect at the outset of the experiment that evidence might be discovered of a function at the judgment level paralleling the appearance of anticipatory lid responses with practice, but a careful analysis of the data has revealed no such concomitance between conditioned lid reactions and changes in J_2. Indeed, there was no very consistent change in these judgments with practice, which is, of course, the basic reason for considering this possible source of modification of J_2 relatively unimportant in the present experiment. It is conceivable, however, that a reinforcing effect of expectancy might prove to be important in explaining the negative time error at longer intervals, as well as the frequently observed changes in the error with practice (**14**, **22**, **33**, **34**). This last suggestion arises out of the fact that the latency of the conditioned response tends to change with practice. The reinforcing effect of expectancy on judgment of the second stimulus of a pair might follow a similar course.

Even though it has not been demonstrated in our experiments that J_2 is affected by establishing sets through practice, it is possible to look for evidence that J_1 has been so affected, particularly in view of the difference observed between J_1 paired and single. It is certainly true phenomenally that if S_1 is usually followed by S_2, the omission of S_2 leaves a sort of blank. There is an experienced reaction to the failure for expectancy to be fulfilled. Actually, however, there is no evidence that J_1 changes in any consistent direction with practice or in a different direction from single J's, though it

is important that the present study be supplemented by another, in which series of single stimuli are presented, in order to observe any change which may occur in judgments under this condition, as compared to the single J's within a series of paired stimuli. It is possible that the set of expectancy is built up too quickly to show in the curves of the present experiment, since the instructions given at the outset of the experiment gave notice that two stimuli would be presented.

The effect of sets resulting from interpolated stimuli, series effects, and more remote levels of judgment will be discussed below as background factors for S_1 and S_2.

III. Reciprocal Interaction

a. Contrast.—In the foregoing discussion, interaction has been assumed to proceed from S_1 to J_2 without any reference to an effect in the opposite direction. There are theories which assume reciprocal effects of S_1 and S_2 processes, in which one gains at the expense of the other. When reinforcement proceeds from the weaker to the stronger of the two processes, so that the weaker is decreased in value and the stronger enhanced, the result is called contrast.[5] In the time error literature, the term 'contrast' seems never to have been applied to the interaction between the two stimuli of a pair being compared, for theories such as Pratt's (**27**) and Needham's (**24**) have emphasized contrast between a stimulus and a background provided by previous experience, by the present series, or by an interpolated stimulus, while Fernberger (**6**) has observed contrast between pairs of stimuli, resulting in a tendency for a light variable following a heavy one, to be judged heavy more often than when it followed a light variable. There are, nevertheless, other theories which have posited interaction and contrast between successive processes in the organism and these should be mentioned even though

[5] It is not clear that all writers have insisted on this reciprocal effect in defining contrast. Many times the word seems to refer to a one way effect. Perhaps either use of the term could be justified. The meaning intended should, in any case, be made definite.

they do not refer specifically to the time order error in their original form.

McDougall's drainage theory (**19**), for example, assumes that when stimuli set up simultaneous or successive processes, drainage may take place from the weaker to the stronger process, so that the stronger is reinforced at the expense of the weaker. In analogies borrowed from hydraulic systems, he has pointed out that the larger and more rapid of two connected streams of water will draw to itself water from the weaker stream. This same theory has been applied to the conditioning situation. Here, again, it is assumed that the greater activity of the center controlling the unconditioned response draws to it the lesser activity set up by the conditioned stimulus. In this fashion it is supposed that a neural connection is made between the conditioned stimulus and the unconditioned response. Razran (**29**) has elaborated this theory, assuming that the more dominant of two S–O–R patterns is reinforced at the expense of the less dominant.[6]

Heymans' law (**11**) is founded on data which might also seem to suggest a contrast effect between stimuli, although only the inhibiting effect of the super-threshold stimulus on the near threshold stimulus is reported and no evidence of enhancement of the larger stimulus processes is given.

Evidence of reciprocal contrast effects is seen in the work of Graham and Granit (**10**, p. 853), which shows that the fusion frequency of the darker of two adjacent patches is decreased by the presence of a brighter patch, while the fusion frequency of the lighter is increased. These authors do not posit a drainage effect, although the result is descriptively called contrast. They assume rather that inhibition increases with increase in intensity so that the brighter spot sends out more inhibitory effect than the darker spot. Fry and Bartley (**7**), Teplov (**30**), Beitel (**1**) and others have made similar observations for certain ranges of intensity relationships.[7]

[6] The experimental evidence for this conclusion is unsatisfactory (**12**, p. 426), as it is also for a drainage theory at the neurological level.

[7] The fact that these experiments all involve spatial separation of paired stimuli introduces an important difference from the present experiment. This does not, however, negate the interest of the comparison.

As far as the present experiment is concerned, there is no evidence for such contrast effect between members of a pair of stimuli. The J_2's are all definitely decreased in value by pairing in spite of the fact that S_2 is never weaker than S_1. Contrast in this case should have resulted in J_2's being enhanced or unchanged. Again, J_1 was increased by pairing where it should have been decreased on a contrast principle. Finally, there were no significantly negative correlations between J_1 and J_2 which one would expect if contrast were operating. If contrast is present, therefore, its effects are obscured by opposing effects.

b. Equilibratory effects.—This interaction effect, like contrast, involves increase in one process at the expense of another, but, unlike contrast, produces enhancement of the smaller of two processes and a corresponding decrease in the larger. In this case, the effect of pairing would be a decrease in the difference between the two processes and a tendency toward equilibrium. The notion suggests Gestalt theories which assume that reactions between fields of different 'concentration' proceed from the direction of the less concentrated field to the more concentrated. The principle has been illustrated in the writings of Köhler, Koffka, and others by physical Gestalten, phi phenomena, and other situations including relative judgments of paired auditory stimuli (**13**, p. 441 ff.). It should be noticed, however, that the Gestalt theories of the time error have not suggested that the weaker stimulus process is enhanced by the stronger, for the typical time error experiments have not obtained independent judgments of two stimuli, and the explanations have been directed only toward the determination of the direction of the relative judgment of 'greater' or 'less.' It is in explaining direction, then, that the Gestalt theories have assumed a leap of electrical potential between two areas of activity set up by unequal stimuli or between the trace of one stimulus and the excitation from another.

There is, in the general experimental literature, very little additional evidence of facilitation effects proceeding from the stronger to the weaker of two stimulus processes at the

expense of the stronger, though there are, of course, many experiments which report facilitation of weak stimuli by strong and the lowering of thresholds by concurrent super-threshold stimulation (**10**, p. 846 ff.; **9**). That this involves inhibition of the stronger process is not stated, however, and the effect is presumably, though not necessarily, a one way affair.

Though parts of the present data are inconsistent with an equilibration theory, it must be recognized that the following facts do appear to be compatible with some such principle: (1) Pairing of stimuli results in the enhancement of J_1 when S_1 has intensity values 1, 2, or 3. This enhancement of the weaker stimulus seems to be accompanied by a decrement in average J_2 values at these same intensities. (2) Subject W who showed less decrement in J_2 also showed less enhancement of J_1. This again suggests inverse concomitant variation in the two judgments. (3) There is a positive correlation between J_1 and J_2 at intensities 3 and 4. This would be expected if an equilibration principle were at work, but the absence of correlation at intensities 1 and 2 would be difficult to explain unless it were assumed that the reciprocal effect was greatest between fields of activity most nearly alike, which is the reverse of the usual assumption that the greater potential difference between fields causes greater activity.

On the other hand, the following points argue against gain in J_1 at the expense of J_2: (1) As already mentioned, J_2 is inhibited at all intensities, *including the situation where* S_1 *has a value of* 4. Decrement in this last case cannot be explained by equilibratory principles, since with an S_1 of intensity 4, J_1 also shows a decrement. (2) Subject E showed marked enhancement of J_1 at the 177 ms. interval, although she showed no decrement in J_2. (3) There is no quantitative relation between the amount of enhancement in J_1 and the amount of decrement in J_2. For example, subjects B, S, E, and D showed greater decrement in J_2 at 285 ms. interval, when S_1 had the value of 1 or 2, than at the shorter interval. In spite of this, they all showed less enhancement of J_1 at 285 ms. than at 177 ms. This again does not support the theory. (4) As pointed out above, if a difference in degree of

activity set up by S_1 and S_2 determines the amount of their interaction, it would be expected that the enhancement of J_1 and the decrement of J_2 would increase as the distance between intensities S_1 and S_2 increased. This evidently does not occur, and again, the result is incompatible with the principle in question. (5) As already indicated, absence of positive correlation at intensities 1 and 2 is difficult to explain. (6) Finally, there are no consistent reciprocal changes in J_1 and J_2 with practice to indicate that one gains at the expense of the other.

It is clearly impossible, on the basis of these facts, either to establish or disprove the operation of a principle of equilibrium. It is certain, however, that if the principle is at work, it cannot account entirely for the presence of inhibition of J_2, inasmuch as this occurs in the absence of reciprocal effects between J_1 and J_2. On the other hand, evidence on the point is suggestive enough to justify further experimental search for critical data, since at the judgment level some such principle as this might work along with the principles of inhibition.

IV. Effects of S_2 on J_1

The preceding paragraphs have been concerned with reciprocal interactions between the after-effects of S_1 and S_2, the assumption being in each case that the one process gains at the expense of the other. It should be pointed out at this time that it is also possible to assume one-way effects of S_2 on J_1, just as we stated earlier that certain after-effects of S_1 might modify J_2. Indeed, this sort of theory is a logical supplement to the theory that S_1 leaves effects which modify J_2. In this case, the effects of S_1 and S_2 would not necessarily be reciprocal, for the influence of S_1 on J_2 might involve different processes and functions from those responsible for the effect of S_2 on J_1. That S_2 should affect J_1 is possible, in spite of the fact that S_2 follows S_1 at an interval of either 177 ms. or 285 ms., because the judgments of the paired stimuli were not *delivered* until after the occurrence of S_2. It is also possible that they were not *made* until after the presentation

of S_2. If it is assumed, as it must be, that some trace of S_1 has persisted during this time, this trace might well be reinforced by delay processes set up by S_2, in much the same manner as suggested in the discussion of facilitation, or there might be direct spread from the S_2 process to the S_1 process without loss to the former.[8] It is not possible to make a clear-cut case for this theory, for S_2 has not been varied systematically as was S_1, and so J_1 cannot be plotted as a function of S_2. It may be significant to notice, however, that there is no consistent difference in the amount of enhancement of J_1 from one intensity of S_1 to another. Since the intensity of S_2 was the same throughout, this would be compatible with a reinforcing effect from S_2 to J_1.

V. Interaction of Stimulus Traces With Background Processes

In the foregoing discussion, no account has been taken of the possible effects of general sets and backgrounds upon successive comparisons, except insofar as the first stimulus itself established such a set. Many writers have suggested that the time order error of stimuli in the series being judged is dependent chiefly on the intensity level of stimuli in previous experimental series (or in non-experimental experience) or of backgrounds interpolated between the pair of stimuli being judged. Although these factors unquestionably play a part under some circumstances, they will be shown not to account adequately for the present findings. Since it is impossible at this time to discuss all the theoretically possible ways in which background factors might affect judgments, only an outline of the most obvious alternatives can be given.

In order for a theory of background effects to be specific, some decision must be made about the following points when absolute judgments of both paired stimuli are involved: (1) Is the trace of the first stimulus the only one affected by background factors, as is usually assumed (**33**), or is the trace of the second stimulus also involved? (2) In either case, does

[8] As suggested in the second paper (**26**) a similar mechanism might be responsible for backward association.

the stimulus trace or traces react with the background by assimilating [9] to it or by contrasting with it? (3) It is then a question whether absolute judgment of the first stimulus grows directly out of the level of the processes mentioned above or whether this judgment is modified further by the processes set up by the second stimulus. (4) If S_2 does interact with the processes discussed in points (1) and (2), does the interaction involve assimilation, contrast, or a one way effect? (5) From what combination of processes does the judgment of S_2 originate? If it is not affected by the background directly, how does it interact with the composite effects of S_1? (6) If it is directly affected by background processes, again the question of the nature of this interaction arises.

Though an attempt has been made to deduce the expected outcome of many of the different combinations of the above alternatives, the extensiveness of the material makes its presentation at this time impracticable. It may be said in general that none of the alternatives seems to fit the facts of the present experiment. At this time detailed consideration will be given to only two of the possible combinations of background factors which are based in part on theories already proposed.

Woodrow (**33**) has developed the assimilation notion in considerable detail and has suggested that two sorts of background may affect the trace of the first stimulus, the series average, and the more remote indifference interval set up by previous experience. The trace of the first stimulus 'gravitates' toward these background effects and the resulting composite trace, which he calls a 'set,' tends to drop off in effectiveness with the passage of time. The composite trace, remaining at the moment of presentation of the second stimulus, furnishes the level against which the process resulting from S_2 is compared. The positive time error will occur, therefore, when the background effect is sufficiently greater than the first stimulus to cause it to 'gravitate' to a level

[9] 'Assimilation' seems to have about the same meaning as 'equilibration' above, though it is not certain that all writers using the term imply a reciprocal effect.

which will be above the second excitation at the moment of comparison. He would explain Köhler's positive errors by the fact that his (Köhler's) variables were all smaller than the standard.

Lauenstein's (**16**) theory was concerned chiefly with interpolated backgrounds and his assumption was that assimilation took place between the trace of the first stimulus and the intensive level of the stimulus interpolated between the stimuli being compared. His theory differs from Woodrow chiefly in its physiological cast. Furthermore, Lauenstein stated specifically that he was unable to obtain evidence of assimilation to interpolated backgrounds at short intervals between first and second stimuli, and this resulted in his acknowledgment that he saw no appropriate explanation of the positive time error.

Since these theories were not suggested with absolute judgment of paired stimuli in mind, it is obviously impossible in the present situation to test them, as usually stated. It is possible, however, to assume that either assimilation or contrast is the principle at work at each step and to predict the necessary result. If assimilation is taken as the principle and an assumption made about each of the questions outlined above, it will be assumed first, that the trace of S_1 assimilates to the general background to form a composite level. This level might be taken as the determinant of J_1, but the present data suggest strongly that J_1 is not independent of S_2. If then, the composite of S_1 and background is to assimilate with the process set up by S_2 before J_1 occurs, J_1 will be affected by S_2 and J_2 by S_1. Under these circumstances, if the average level of our series is taken to lie between stimuli 3 and 4,[10] it should follow that when S_1 has a value of 4 and S_2 is equal to it, the S_1 process will 'gravitate' down toward the background, the composite falling between background and the intensity 4 level. The S_2 process would then gravi-

[10] This appears to be a justifiable assumption, for although three weaker intensities lay below intensity 4, this intensity occurred 22 times to 6 of each of the other intensities. Also, it must be remembered that the distance between intensity 3 and 4 is approximately 49 times as great as that between 2 and 3.

tate down toward the composite and the composite up toward S_2. If it is assumed, as most theories have done, that J_2 is not directly affected by the background but only by this interaction with the S_1 composite, then J_1 should be less than J_2, though in fact it is judged greater. Furthermore, the observed relation between amount of decrement of J_2 and intensity of S_1 would not be expected. The weakest S_1 should yield the smallest J_2, by drawing down the S_2 trace or leaving it unrelated to the size of S_1, if the S_2 process were not assumed to 'gravitate.' The observed increase in J_1 paired would be expected for all intensities of S_1 used here, provided the second gravitation step of the composite to S_2 process were kept as part of the theory. If this were omitted, as it is in Woodrow's theory, it would be necessary to assume that only paired stimuli assimilated to the background, in order to explain pairing effects on J_1.

When an elaborated contrast theory is subjected to the same examination, it appears to have even fewer advantages. In this case, the same assumptions are made as above, except that where assimilation was posited, contrast is now assumed to be at work. Such a theory is capable of explaining the positive time error at the largest intensity of the first stimulus, for the trace of S_1, which must be greater than the series average, will be raised and its size exaggerated thereby. Since the processes set up by S_2 are not assumed to contrast directly with the background, they will be smaller than the S_1 composite and contrast between the composite and S_2 will further exaggerate the difference, causing J_2 to be much smaller than J_1. The increase in J_2 as the intensity of S_1 decreases is also to be expected on this theory, though it does not explain the terminal inversion of J_2 when S_1 has the value of 4. The enhancement of J_1 paired is completely inconsistent with the results that would be predicted from the theory since an S_2 greater than S_1 should decrease, rather than increase J_1, by contrasting with it. Finally, negative instead of positive correlations between J_1 and J_2 would be expected.

Summary

1. The experiment upon which the present discussion is based has demonstrated the presence of a positive time error at intervals of 177 ms. and 285 ms. between stimuli and has revealed the fact that the 'error' is due both to a falling off in the apparent loudness of the second stimulus and to an enhancement of the judged loudness of the first as a result of pairing.

2. The paper has classified the sources of modification of judgments into five groups and has discussed the applicability of these factors to the details of the experimental findings.

3. Certain generalizations regarding the observed inhibition of responses to the second stimulus of a pair in spinal reflexes, in the lid reflex, and in judgments of paired tones, have been offered as a possible part of a theory of the positive time order error. These findings are thought to suggest the possibility that similar determinants are at work in producing the functions in question. When, therefore, the factors causing inhibition in one of these reaction systems are discovered, they should be considered as possible determinants of the other similar functions and of the positive time error. The comparison of behavior of different kinds may thus be the source of fruitful hypotheses regarding the general principles of behavior, and through experimental test of such hypotheses the events correlated with the positive time error may be discovered. Such correlates will of course constitute part of the explanation of the error.

4. These generalizations regarding the inhibition of J_2 do not, however, throw any light on the enhancement of J_1 by pairing which must also be accounted for in a theory of the positive time order error. A number of theories relating to this point have been discussed and the following appear to afford the most promising possibilities:

a. An equilibration theory gains some support from the fact that in certain cases J_1 seems to be enhanced at the expense of J_2. There are, however, serious exceptions to this relationship that weaken its plausibility, and it cannot possibly

account for the inhibition of J_2 at the largest intensity of S_1 where J_1 is not enhanced but is slightly decreased by pairing.

b. Theories of the facilitation of J_1 by long-circuited processes set up by S_2 are suggested as the logical supplements of the inhibition notion, though the present experiment does not afford any definite data on the functional characteristics of such facilitation if it occurs.

5. Finally, it is suggested that negative time errors may prove to be related functionally to facilitation processes arising from S_1 at longer intervals and that comparisons of such functions with facilitation observed at other behavior levels might prove fruitful in revealing further similarities and differences in reactions of the organism.

REFERENCES

1. Beitel, R. J. Human retinal synaptic functions, summation and inhibition. *Amer. J. Optom.*, 1936, **13**, 452–454.
2. Borak, J. Über die Empfindlichkeit für Gewichtsunterschiede bei abnehmende Reizstarke. *Psychol. Forsch.*, 1922, **1**, 374–389.
3. Creed, R. S. *et al. Reflex action of the spinal cord.* Oxford: Clarendon Press, 1932.
4. Culler, E. Studies in psychometric theory. I. The Fechnerian time-space errors. *Psychol. Monogr.*, 1926, **35**, Whole no. 163.
5. Dusser de Barenne, J. J. G., & McCulloch, W. S. Local stimulatory inactivation within the cerebral cortex. *Amer. J. Physiol.*, 1937, **118**, 510–524.
6. Fernberger, S. W. Interdependence of judgments within the series for the method of constant stimuli. *J. exper. Psychol.*, 1920, **3**, 126–150.
7. Fry, G. A., & Bartley, S. H. The effect of steady stimulation of one part of the retina upon critical flicker frequency in another. *J. exper. Psychol.*, 1936, **19**, 351–356.
8. Fulton, J. F. *Physiology of the nervous system.* Oxford Press, 1938.
9. Geldard, F. A. Brightness contrast and Heymans' Law. *J. gen. Psychol.*, 1931, **5**, 191–206.
10. Graham, C. H. Vision. III. Some neural correlations. In *A Handbook of General Psychology*, Worcester: Clark University Press, 1934.
11. Heymans, G. *Gesamelte kleinere Schriften Untersuchungen über psychische Hemmung.* Haag: Nijhoff, 1927, Part II.
12. Hull, C. L. Learning. II. The factor of the conditioned reflex. In *A Handbook of General Experimental Psychology*, Worcester: Clark Univ. Press, 1934.
13. Koffka, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace, 1935.
14. Köhler, W. Zür Theorie des Sukzessivvergleichs und der Zeitfehler. *Psychol. Forsch.*, 1923, **4**, 115–175.
15. Kreezer, G. The neurological level of the factors underlying time errors. *Amer. J. Psychol.*, 1938, **51**, 18–43.
16. Lauenstein, O. Ansatz zu einer physiologischen Theorie des Vergleichs und der Zeitfehler. *Psychol. Forsch.*, 1932, 130–177.

17. LEHMANN, A. Beiträge für Psychodynamik der Gewichtsempfindungen. *Arch. f.d. ges. Psychol.*, 1906, **6**, 425.
18. LORENTO DE NÓ, R. Inhibition of motoneurones. Reprinted from *Problems of Neural Physiology and Behavior.*, 1936.
19. McDOUGALL, W. The bearing of Professor Pavlow's work on the problem of inhibition. *J. gen. Psychol*, 1929, **2**, 231–262.
20. MÜLLER, G. E., & Martin, L. J. Zür Analyse der Unterschiedsempfindlichkeit. *Leipzig*, 1899. Quoted from (**23**).
21. MÜLLER, G. E., & SCHUMANN, E. Über die physiologische Grundlagen des Vergleichen gehobenen Gewichte. *Arch. f.d. ges. Physiol.*, 1899, **45**, 37 (not seen).
22. NEEDHAM, J. G. The time error as a function of continued experimentation. *Amer. J. Psychol.*, 1934, **46**, 558–567.
23. ——. The time error in comparison judgments. *Psychol. Bull.*, 1934, **31**, 229–243.
24. ——. Contrast effects in judgments of auditory intensities. *J. exper. Psychol.*, 1935, **18**, 214–226.
25. PEAK, H. Time order errors in successive judgments and in reflexes. I. Inhibition of the judgment and the reflex. *J. exper. Psychol.*, 1939 (in press).
26. ——. Time order errors in successive judgments and in reflexes. II. As a function of the first stimulus of a pair. *J. exper. Psychol.*, 1940 (in press).
27. PRATT, C. C. Time errors in the method of single stimuli. *J. exper. Psychol.*, 1933, **16**, 798–814.
28. ——. The time error in psychophysical judgments. *Amer. J. Psychol.*, 1933, **45**, 292–297.
29. RAZRAN, G. H. S. Theory of conditioning and related phenomena. PSYCHOL. REV., 1930, **37**, 25–43.
30. TEPLOV, T. M. Dependence of the absolute visual threshold on the presence of an additional stimulus in the field of vision. *J. gen. Psychol.*, 1936, **15**, 3–11.
31. WADA, Y. Der Zeitfehler beim Sukzessivvergleichs der Tonhöhen. *Jap. J. Psychol.*, 1932, **7**, 505 (not read).
32. WHEELER, R. H., & PERKINS, F. T. *Principles of mental development.* New York: T. Y. Crowell Co., 1932.
33. WOODROW, H. The effects of practice upon time order errors in the comparison of temporal intervals. PSYCHOL. REV., 1935, **42**, 137–153.
34. ——, & STOTT, L. The effect of practice on positive time order errors. *J. exper. Psychol.*, 1936, **19**, 694–705.

[MS. received July 6, 1939]

AN EXAMINATION OF THE CONCEPTS OF DOMINATION AND INTEGRATION IN RELATION TO DOMINANCE AND ASCENDANCE [1]

BY HAROLD H. ANDERSON

University of Illinois

In much of the psychological literature the terms *domination*, *dominance*, and *ascendance* are used interchangeably. The writer has previously offered a definition together with illustrations of the terms *dominate*, *domination*, and *dominating* (**5**, **6**), which three terms are used by him to designate similar psychological phenomena. In general this definition is in agreement with the use of these terms elsewhere. There is no such consistency, however, to be found in the literature in the use of the terms *dominant* and *dominance*. It is suggested here that the inconsistencies to be found in the use of the terms which are included in the scope of this discussion are due to the use of the same label for phenomena that are psychologically different. The intention here is not to quibble over the use of terms. The purpose is merely to attempt a statement which will lead toward a better understanding of certain psychological concepts and a further clarification of their meanings.

Domination.—In the following quotation as well as in other publications (**3**, **4**, **5**, **6**, **9**, **10**) the writer has defined domination as a technique of behaving, as a form or mode of responding to one's environment. In each of the above citations the term *integration* has also been defined.

"The use of force, commands, threats, shame, blame, attacks against the personal status of an individual are called dominative techniques of responding to others. Domination

[1] This paper was presented May 6, 1939, at the symposium on "Principles of Child Development," Midwestern Psychological Association Meetings, University of Nebraska, Lincoln.

is characterized by a rigidity or inflexibility of purpose, by an unwillingness to admit the contribution of another's experience, desires, purposes or judgment in one's determining of goals which concern others." "Domination is an attempt at atomistic living; the desires, purposes, standards, values, judgment, welfare of others do not count; it is rugged individualism of a highly ingrowing order. Domination is the antithesis of the scientific attitude; it is an expression of resistance against change; it is consistent with bigotry and with autocracy. It is the technique of a dictatorship" (7, p. 73).

Integration.—The terms *domination* and *integration* are used by the writer merely as convenient labels for two techniques of behaving that have been experimentally demonstrated to be psychologically different. In the writer's initial investigations in this area it was assumed for example in the behavior of preschool children that there is a psychological difference between snatching a toy out of a companion's hands so as to play with it oneself and asking the companion if one may borrow the toy for awhile. It was assumed at all age levels that there was a psychological difference between a command and a request, between "tellin' 'em" and "eskin' 'em."

If, instead of compelling the companion to do as one says, one asks the companion and by explanation makes the request meaningful to the other so that the other can voluntarily coöperate, such behavior is said to be an expression not so much of pursuing one's own unique purposes as attempting to discover and get satisfactions through common purposes. For such expenditure of energy in common purposes, for an attempt to reduce instead of to augment or incite conflict of differences the term integrative behavior is used. The person who can change his mind when confronted with new evidence which has grown out of the experience of another is said to be integrating differences.

The term integration is not used here as it has been used by some in contrast with differentiation. It is believed that the two processes are inseparable and are merely different aspects of the same psychological or biological phenomenon. With the integration of differences something new is created that never has

existed before; this emergence of originals through the integration of differences is itself a differentiation.

Whereas domination stifles or frustrates individual differences, socially integrative behavior respects differences, advances the psychological processes of differentiation. Integrative behavior as the term is used here is consistent with the scientific point of view, the objective approach. It designates behavior that is flexible, growing, learning. It is an expression of the operation of democratic processes (7, pp. 73–74).

Dominance.—The terms *domination* and *integration* have been used above to designate techniques of responding that have been reliably observed and experimentally demonstrated to be different. The terms *dominant* and *dominance* have been used in the literature to designate relationships that at one time are created and maintained by techniques of domination and at other times are based on techniques of a highly integrative nature. Can one be *dominant* without being *dominating?*

A slave driver is *dominant* as compared with his slaves. It is *his* desires that determine the work to be done; it is *his* judgment that determines the way in which the work is to be done. He may crack the whip to enforce his *dominance*; or he may never use the whip but achieve dominance through the force of law or the power of convention.

The dog in the manger, although he could not use and did not want the hay, was *dominant* over the cow.

The chicken that pecks the greatest number of other fowls in the flock is reported by psychological investigators to be *dominant.* The hog that crowds out all-comers from the trough is *dominant.*

But the slave driver, the dog in the manger, the dominant chicken, and the dominant hog are also in the writer's terminology *dominating*: they are in their respective relationships dominant because of their use of force, or threats or other techniques of expending energy *against* others which are the criteria of *domination.* In addition, they appear to be in control of a greater balance of *power over* others. The others, by contrast, being at the low end of a balance of power in a conflict of purposes are said to be submissive. Submission

implies duress or coercion and may be said to characterize behavior in degrees which vary proportionally with the adverse momentary balance of power in a conflict situation.

Psychologists generally use the terms *dominant* and *dominance* to express a relationship between two individuals or between the individual (or organism) and his environment. *Dominant* and *dominance* are, however, rarely distinguished from domination as this technique has been generally defined. The words all come from the same Latin root and in psychology have the connotations given in the dictionary definitions of 'ruling, prevailing, controlling, commanding.' Dominance therefore currently implies a 'superiority, ascendance' that has been achieved by the technique of domination.

If there are differences between the concepts of *dominant* and *dominating*, these differences are not generally regarded by psychologists as being important, as can be seen by examining almost any text on general psychology. Take for example Dashiell's *Fundamentals of General Psychology* (**12**). Dashiell uses the terms *dominant* and *dominating* in introducing his discussion of *ascendant* behavior. Dashiell writes:

> When people meet, what often happens is that certain ones come to play the more dominant rôle while the others play the more submissive. Is this characteristic of dominating (ascendance) or following (submission) one single general and consistent fact about an individual? (**12**, p. 586).

The context from which the above quotation was taken offers no further qualification of the synonymous use of the terms dominant, dominating, and ascendance.

Maslow in the Self-Rating Scale of his Social Personality Inventory (Short Form for Women) gives the following explanatory instructions:

> You are to rate yourself on the feeling of dominance in yourself. Feeling of dominance is defined as general self-confidence, and high level of self-evaluation. It includes forcefulness of character, 'strong' personality, a feeling of being 'on top,' a feeling of mastery over others, a feeling of general capability. It may assume several forms in different people: despotism, aggressiveness, helpful leader-

ship, guidance, or just feeling dominant with no active expression. Yet the same feeling of dominance is present no matter what aspect it takes.

Low dominance-feeling, in general, consists of the reverse or absence of these feelings. It includes feelings of general inferiority, lack of self-confidence, uncertainty, feeling dominated by others, timidity and shyness, often a willingness to be self-effacing, and self-consciousness.[2]

Maslow recognizes that "a feeling of dominance . . . may assume several forms. . . ". For example, 'mastery over others,' or 'aggressiveness, an inclination to push oneself forward, imposing one's wishes on others' would satisfy the criteria for domination, for the expenditure of energy against another. These phrases could well be used to classify the extreme dog-in-the-manger behavior. 'Helpful leadership, guidance,' which are listed by Maslow immediately after the terms 'despotism, aggressiveness,' may be assumed to be different forms of what Maslow construes as dominance.

Maslow uses the term 'dominate' to designate behavior that includes a range of relationships from high degree of conflict in purposes to a high degree of commonness of purpose.

The fact that *dominance* when used synonymously with *domination*, as defined for the present discussion, is inadequate to describe Maslow's use of the term *dominance* is specifically mentioned by Maslow. In discussing "Dominance status and dominance-feeling" Maslow writes:

Dominance status.—It may be readily observed in pairs of human beings that, in a good many cases, one member of the pair will dominate the other. He will lead, suggest, order or inhibit the behavior of the subordinate one. The latter will often defer to the wishes of the dominant one, and respect him more than the dominant one respects *him*. In a marriage, for instance, in which the wife definitely and strongly dominates the husband, her opinion will usually prevail whenever there is any conflict; she will be more forceful and direct in the expression of her opinions than he will. In their social relationship, she is more confident in her behavior,

[2] Quoted from Eisenberg (13, p. 16).

gets what she wants more often than does the husband, feels superior to him, generally feels herself to be stronger than he is, and respects herself more than she does him. Her behavior, depending on the kind of dominance, is apt to express either (1) condescension, contempt, impatience, irritability, (2) pity, tenderness, protectiveness or good-natured toleration, or (3) aloofness and superiority (**16**, p. 405).

To the present writer and probably also to Maslow the behavior listed in (1) and (3) in the above quotation represents a conflict of differences and the expenditure of energy resisting differences in others. On the other hand the behavior listed in (2) as 'pity, tenderness, protectiveness or good-natured toleration' probably represents a community of interests and a rapprochement of differences. The techniques which one might expect to observe in (1) and (3) would be techniques of domination; in (2), however, one would expect the behavior to be highly flexible, integrative.

The assumption that the confronting of differences need not represent a conflict of differences underlies the concept of integrative and coöperative techniques of behaving; it is also basic to the concepts referred to by psychologists as 'objectivity' or a 'scientific attitude.' That two persons can live on the face of the earth and not for any inherent, instinctive or otherwise 'necessary' reason be in conflict is implied by Maslow. Discussing *dominance status* and *subordinate status* Maslow says, "While this kind of relationship does not *always* obtain in social pairings, it is very common in our culture" (**16**, p. 406).

In discussing his concept of 'dominance-feeling' Maslow explains certain difficulties and ambiguities in defining the concept. By way of emphasis after listing "the words and phrases which dominant people themselves have used when asked to describe their own feelings about themselves," Maslow adds in a footnote:

It is necessary at this point to insist as strenuously as possible that dominance-feeling be not confused with domineeringness, with antagonism to others, willfulness, impoliteness, selfishness, aggres-

siveness, tyrannizing, etc. If the reader finds himself prone to think of a high-dominance person as rather a nasty person, one who would push him out of line or punch him in the face, let him substitute 'self-confidence' for 'high-dominance feeling,' and 'lack of confidence' for 'low-dominance feeling.' They are not synonymous, but at least the substitutes bring less misunderstanding in this direction. The writer uses the word 'dominance' because, even after much thought, he has not been able to find one that is not also misunderstood, and at least our phrase makes the animal work continuous with work for children and adult human beings. Another possible substitute is high or low ego-level (**16**, p. 408).

The phrases which dominant people have used in describing their feelings to Maslow are ambiguous as to whether they imply techniques of domination or even 'domineering' as, for example, '(8) a feeling of being able to handle other people, (9) a feeling of masterfulness and of mastery' (**16**, p. 407). Either of these phrases *could* include domineeringness and domination. Also, 'self-confidence' as used in interviews is a term that could be construed as describing at one extreme the dominance-feeling of a brutal father about to strap a child or the dominance-feeling of a minister with a good sermon all prepared who is mounting the rostrum to address his loving and coöperative flock.

Eisenberg, who, for his concepts of dominance, acknowledges indebtedness to Maslow (**13**, p. 6) has been troubled in the theoretical interpretations of his investigations by apparent inconsistencies in human and primate behavior as explained in terms of dominance.

Though we cannot describe them, there seems a possibility that within dominance and non-dominance as such sub-syndromes may be described. . . . Roughly these may be described as: (1) the 'tyrant' or despot who demands all of the person he dominates; (2) the 'coöperator,' the one who controls the other but with the socially desirable purpose of facilitating the smooth performance of a task; (3) the 'helpful guide' or the 'teacher,' the individual to whom others come for advice or aid (this may be very much the same as the coöperator); (4) the individual who makes no attempt to control others, yet feels self-confident and permits no one to control him (he is very likely not very sociable) (**13**, p. 67).

It is clear that the techniques of Eisenberg's 'tyrant' or despot are techniques of domination. The techniques of the 'coöperator' and 'helpful guide' are more probably techniques of integrative behavior as this latter term is used by the writer. If there is a neutral or non-functioning relation between dominative behavior on the one hand and integrative behavior on the other it would probably be found in the non-confronting of differences indicated by Eisenberg's fourth classification which includes 'the individual who makes no attempt to control others' and who 'is very likely not very sociable.'

An example of another experimental study where a need is expressed for more precision of meanings in the use of the term *dominant* is found in a short report by Hanfmann. Hanfmann defines what she means by dominant, but at the same time recognizes that the term does not adequately describe or differentiate the behavior which she observed as being dominant. Dominance, she found, did not always mean the same thing.

Hanfmann's definition of dominant satisfies also the criteria for domination; in fact she uses the two terms interchangeably. However, the fact that Hanfmann failed to observe in her kindergarten subjects a linear relationship or hierarchy of dominance led her to raise some questions.

> . . . While A dominates B, and B dominates C and D, each of these: C and D in his turn dominates A. This circular connection suggests that the dominance does not mean the same thing in all cases, and actually we find that these four or five children represent different patterns and different methods of dominance (14, p. 407).

Hanfmann's difficulty with the conventional use of terms led her to describe as *negative dominance* the behavior of one child who resisted attempts of other children to bring about joint play.

Of the four children, A, B, C, and D, who according to Hanfmann rank highest in 'dominance,' two, A and B, show behavior that in the writer's terminology is definitely *domi-*

nating and two, C and D, show behavior that is socially *integrative.* The distinction which the writer makes between commands (domination) and requests or suggestions (integrative behavior) is not made by Hanfmann expect that the 'domination' found in child C and child D (which the writer calls not domination, but integrative behavior) is referred to by Hanfmann (and by the writer) as *leadership.*

Hanfmann's detailed description of the behavior of child D to whom she refers as a *social leader* is almost point for point an illustration of growth behavior by a secure individual who accepts others as they are, who seeks and finds a common purpose among differences, who increases the understanding between himself and others, who makes others secure in his presence; in no particular does this 'dominant' child 'dominate' his companion.

> . . . child D . . . tries to induce and to maintain the joint play whenever he can, and actually succeeds with six children out of nine. His methods of controlling the other child are similar to those of C, for he bases his demands upon the requirements of the play, never coercing his playmate and never flatly refusing his requests. But he goes further than that: he actually takes into account the attitude of his companion and displays an astonishing skill and flexibility in trying to remove rather than break his resistance: He introduces mutual play not by taking the other child's blocks but by offering his own; he takes care to announce and to explain his actions in advance, especially when taking the blocks away; while the blocks are in his possession, he keeps telling the other one reassuringly that he will return them soon, and actually does so; if he meets with resistance, he never presses the point, and lets some time pass before repeating his request, usually in a modified form; occasionally he simply ignores a refusal or, by treating it as a joke, makes the other child treat it in the same way. Usually these varied diplomatic methods are successful but, in the case of a serious conflict, he will give way, at least temporarily, to his comrade's plans, rather than give up the joint play. We may speak of this child as a *social leader.*

For child B social play has a totally different function—it is a means of expanding and displaying his own power. Though extremely active, he is less interested in the play itself than in

getting all blocks for himself, in being the initiator of all activities, and in having an unrestricted command over the other child. . . . The methods he uses . . . grabbing blocks violently, disregarding protestations, denying the other's requests by saying merely "No, you can't. They are mine. I will do it myself"; suppressing the other's attempts at activity by force, for instance by holding his hands; . . . . In short, the methods used . . . are characterized by great forcefulness and by complete disregard for the rights and wishes of the other child (14, p. 409).

Child A "deliberately disrupts the situation of orderly play. The method of child A is destruction. He himself does not build, and when the other child starts building, he purposely destroys his work, throwing the blocks to the floor, and meeting protests merely with outbursts of laughter. . ." (14, p. 409–410).

Hanfmann found that her subjects, the kindergarten children themselves, made distinctions similar to those discovered in her own psychological observations. Upon being questioned by the teacher the children were not only able to differentiate *dominant by force and coercion* from *dominant by kindness, coöperativeness and understanding;* the 'dominated' children actually preferred the latter form of 'dominance.'

To say of a relationship that one individual is *dominant without being dominating* is to court ambiguity and confusion. And yet psychologists have in their writings, in the descriptions of their observations, reported situations in which one individual satisfies the criteria of dominance or dominant in a relationship which is voluntarily accorded by others through the technique of integrative behavior rather than through the imposition of force or coercion.

Dominance and *dominant* in such cases designate a superior relationship with no implication as to the technique by which it was achieved or maintained. When the technique of achieving dominance is one of domination the case is relatively simple. When the technique of interplay is a democratic, coöperative, or integrative technique, whereby one is accorded by his fellows a relative position of dominance, the issue is not so clear. A main difficulty in the second case is the

arbitrariness of any comparison and the difficulty, if not impossibility, of arriving at a clear concept of the basis of comparison.

To take two illustrations: The case of the dog in the manger is a case of dominance through dominating behavior. One is dominant not so much through the interplay of differences as through the suppression of the interplay of differences. The desires, needs, purposes, motives, experience, or whatever it is that gives the dog his uniqueness in comparison with the cow are the basis out of which is determined the behavior of the dog and the behavior of the cow. The cow presumably would behave differently except for the domination (the techniques, the use of force, threats) of the dog. The dog is not a leader but a driver, a coercer.

As a second illustration take a man and his doctor. The man feels ill. He calls his doctor. His doctor does not come without being invited and he makes no diagnosis contrary to his own unique experience or judgment. The man tells all he can about his symptoms. The doctor tells the man that he believes it is a case of appendicitis and recommends an immediate operation. The man says he can walk to a taxi, goes to the hospital, does as he is told by doctor and nurses. Here is the confronting of individual differences. The doctor is not coerced, neither is the man. The doctor respected the man's judgment. The man respected the judgment, had confidence in the skill of the doctor. The relationship was coöperative in a common purpose of mutual advantage. The man gets an appendectomy, the doctor gets his fee. Who was dominant? Was the doctor dominant? He was if the base line was defined as one involving medical and surgical skill, but only if the comparison is so defined. When the man called the doctor on the telephone who was dominant?

The question may well be raised whether the concept of dominance apart from the technique of domination is not descriptively meaningless.

Leadership.—It seems almost impossible to discuss *dominance* without becoming involved in a concept of *leadership.*

It might be said that if one could be *dominant* without

being *dominating*, he would become a *leader*. In the democratic or integrative processes dominance becomes a highly arbitrary and evanescent concept.

To accept the principle or the fact of individual differences is to necessitate one's admitting the coarseness of all comparison. It also necessitates one's admitting that, irrespective of the process of evaluation of these differences, each person can make a contribution that no one else can make. If one is then to make comparisons he must say that from moment to moment or from situation to situation or in certain arbitrarily defined respects each person in a truly democratic group is at some time because of his uniqueness making a 'superior' contribution to the others.

In a true democracy every individual would become at some time a leader. The concept of leadership would lose the fictitious prestige or superiority which it holds in an authoritarian, competitive culture that only in a very gross way recognizes and acts upon the principle of individual differences. Leadership today possesses a spurious glamor which implies domination or the use of power over others. Domination should play no part in a psychological concept of leadership. Domination reduces the social interplay of differences, stifles the emergence of originals. A *fuehrer* who needs concentration camps to deal with differences cannot to that extent be a *leader*. He is a driver, a despot. He drives because he cannot lead.

A true leader in a psychological sense is one who can make the most of individual differences, who can bring out the most differences in a group, and therefore reveal to the group a sounder base for defining common purposes. He is as much a follower as a 'leader.' He recognizes merit wherever he finds it, is open to new data, is flexible, responsive.

The *value* of individual contributions becomes not so much a matter of relative similarity as a matter of difference. In attempting a logical or consistent definition of leadership the importance of individual differences becomes acute as does the arbitrariness of standards of comparison. A leader becomes one who can facilitate the social interplay of differences, the

emergence of originals, the defining and re-defining of common purposes, processes which are at the foundation of any concept of growth. If one is dealing with concepts of growth and of originals, comparisons are almost out of order. There are no standards in biology for saying that the heart is leader over the lungs and the spleen or that it makes a 'greater' contribution to the biological processes.

If it were possible to divorce the concept of dominance from the technique of coercion, a leader could be said to be *dominant* only because, through his superior ability to understand persons, to help them to become spontaneous, functioning social-participators, he respects individual differences and in this way he reveals in each member of the group a social value.[3] But when the 'dominant' democratic leader listens to and learns one small bit from the most modest and humble of his followers, it would be as proper to say at that moment that the follower had become dominant. The question may again be asked: What do the terms dominant and dominance contribute to the understanding of human beings or their behavior where there exists a process of coöperative, integrative confronting of differences? What descriptive uses have the terms dominance or dominant beyond an expressed or implied concept of some use of force or coercion?

Ascendant behavior.—The concept of *ascendant* behavior carries an ambiguity in the psychological literature similar to the ambiguity found in the use of the term *dominant.* A citation from Dashiell has been given above to show that *ascendant* is not distinguished from *dominant* or *dominating* but is used synonymously. Dashiell also uses the terms 'dictator,' 'bosses,' 'peck . . . with impunity' to designate *ascendant behavior* (**12**, p. 586). It is unmistakable that Dashiell has in mind 'power over' another, the use of force or coercion, rigidity of response and other characteristics of dominating behavior as basic to ascendant behavior. Having said all these things, Dashiell gives as illustration in his text an

[3] An able discussion of the nature and functions of leadership is presented by T. N. Whitehead in *Leadership in a free society* (**19**), and in *Leadership within industrial organizations* (**18**).

example chosen from the Allport Ascendance-Submission scale (2).

The ascendant person, as described by Dashiell, elevates himself by stepping on the neck of someone else. He becomes superior only because someone else relative to himself is inferior. He is ascendant because someone else is submissive. The scale is called by its originators an Ascendance-Submission scale. These descriptions are also consistent with the writings of Allport, as for example:

> The nature of the struggle for survival in a competitive society tends to force every individual to seek his own most suitable level of *aggression*. As the saying goes, everyone *must* be either a boot or a doormat. One child as he matures finds that a constant effort to dominate his fellows is for him the most successful design for living; another finds that for him there is more satisfaction in a characteristic yielding or submission. Somewhere between the extremes of exaggerated domination and complete passivity, there lies for each normal individual a level of adaptation that fits his intimate requirements. The psychologist does well to recognize all these possible gradations and to postulate a common variable (in this case, ascendance-submission) which, though rough and approximate, permits quantitative scaling (1, p. 298).

Both to Dashiell and to Allport, it would seem that individual differences are conflicting differences, and that in seeking one's own most suitable level of aggression human beings live mainly according to a psychology of pulling the cat's tail. One might almost paraphrase Allport: The nature of the struggle for survival in a competitive society tends to force every individual to seek his own most suitable cat!

The following illustration which Dashiell cites from Allport seems to contain not the slightest relationship to the behavior which he has *described* as being *ascendant* behavior and to indicate that both Dashiell and Allport have missed the psychological point of the illustration.

> 1. At church, a lecture, or an entertainment, if you arrive after a program has commenced and find that there are people standing but also that there are front seats available which might be secured without 'piggishness' or discourtesy, but with considerable conspicuousness, do you take the seats?

Habitually.....
Occasionally.....
Never..... (12, p. 587).

To the writer the above test item is simply a test of the individual's *spontaneity*. It asks: Can you be yourself? Can you use your *own* judgment in this situation as defined? Can you do the thing that according to *your own* desires and *your own* purposes seems to be the most appropriate thing to do? Are you spontaneous? Are you secure? Dare you have ideas of your own? And how often: habitually, occasionally, or never, for this situation as defined?

In this test situation the Allports rule out by definition the characteristics which they consider the most important aspects of ascendant behavior. There is no 'boot and doormat' relation permitted, because one can by definition obtain a seat "without 'piggishness' or discourtesy." By definition one cannot render others insecure; by definition one does not make himself comfortable by making someone else uncomfortable. As the situation is defined the people standing are those who either do not want to sit or who are not convinced of the genuineness of the way in which the situation is defined. The real catch that has probably misled the Allports, Dashiell and others is in the word 'conspicuousness.' The implication of 'considerable conspicuousness' is that it is discourteous or that it attacks the status of others, thereby defining a situation where one *can* be *ascendant over* someone else.

To record ascendant behavior *as defined* by Allport and by Dashiell the quoted paragraph would have to be changed to include boos or hisses to those who come in late or to include discourtesy to the speaker or inconvenience to the audience. As the test item stands it is a clearly stated example of a case where differences may exist without being conflicting differences, where one may be himself and live harmoniously with others, be accepted as he is.

The work of the Allports which resulted in the Ascendance-Submission scale has constituted a valued stimulation to psychological thinking and to further experimentation. Allport's scale has been taken over bodily and incorporated with

material from three other scales into an even more widely used personality scale by Bernreuter (**11**). It was also taken over in substance by Jack (**15**, pp. 7–65), and Jack's techniques were employed by Page (**17**) in experiments on preschool children.

Ascendance as a term characterizing personality does not, however, differentiate between flexibility and inflexibility in the techniques of responding to differences, nor between the use of commands and requests in attaining one's objectives.

According to the Jack definitions of ascendant behavior the following activities of preschool children are recorded as ascendant behavior and are assigned equal score values.

1. A child reaches across the sand table and snatches a toy out of the hand of a companion and plays with it himself.

2. A child asks his companion if he may play with a toy if he will give it back in just a minute.

According to the concepts of the writer, the first example is dominative behavior; the second is integrative behavior. The first is frowned upon in nursery schools; the second is encouraged.

As an incidental part of an investigation of domination and integrative behavior in preschool children (**6**), simultaneous records of ascendant behavior as defined by Jack and dominative and integrative behavior as defined by the writer were made on ninety-eight cases. Ascendant behavior had coefficients of correlation of $.65 \pm .02$, with domination scores, and $.44 \pm .03$, with integrative behavior. In the same study domination was shown to have a zero correlation with a companion's integration score ($r = -.10$ on 1030 pairings) and a zero relation to the child's own integration scores ($r = -.07$ on 1030 pairings).

The writer has thus shown experimentally that domination as defined is psychologically different from integrative behavior and that the concept of ascendant behavior includes both dominative and integrative behavior without distinction.

BIBLIOGRAPHY

1. ALLPORT, G. W. *Personality.* New York: Henry Holt, 1937. Pp. xiv+588.
2. ——, & ALLPORT, F. H. *The A-S reaction study.* Boston: Houghton Mifflin, 1928.
3. ANDERSON, H. H. Conflicts in personality development. *Mental Hygiene,* 1936, No. 4, pp. 605–613.
4. ——. Dynamic nature of personality. The National Elementary Principal, Fifteenth Yearbook, 1936, **25,** No. 6, pp. 244–253.
5. ——. An experimental study of dominative and integrative behavior in children of preschool age. *J. Soc. Psychol.,* 1937, **8,** pp. 335–345.
6. ——. Domination and integration in the social behavior of young children in an experimental play situation. *Genet. Psychol. Monogr.* 1937, **19,** pp. 341–408.
7. ——. The measurement of domination and of socially integrative behavior in teachers' contacts with children. *Child Developm.,* 1939, **10,** pp. 73–89.
8. ——. Domination and social integration in the behavior of kindergarten children and teachers. *Genet. Psychol. Monogr.,* 1939, **21,** pp. 287–385.
9. ——. The construction of a mental hygiene scale for teachers. (Accepted for publication in *Amer. J. Orthopsychiat.*)
10. ——. Domination and social integration in the behavior of kindergarten children in an experimental play situation. *J. Exper. Educ.,* December 1939.
11. BERNREUTER, R. G. Personality inventory. *J. Soc. Psychol.,* 1933, **4,** pp. 387–405.
12. DASHIELL, J. F. *Fundamentals of general psychology.* New York: Houghton Mifflin, 1937. Pp. vii+655.
13. EISENBERG, P. Expressive movements related to feeling of dominance. *Arch. Psychol.,* No. 211, 1937, pp. 73.
14. HANFMANN, E. Social structure of a group of kindergarten children. *Amer. J. Orthopsychiat.,* 1935, **5,** pp. 407–410.
15. JACK, L. M. An experimental study of ascendant behavior in preschool children. In Jack, L. M., and others, Behavior of the preschool child. *University of Iowa Studies: Studies in Child Welfare,* 1934, **9,** No. 3, pp. 7–65.
16. MASLOW, A. H. Dominance-feeling, behavior, and status. PSYCHOL. REV., 1937, **44,** pp. 404–429.
17. PAGE, M. L. The modification of ascendant behavior in preschool children. *University of Iowa Studies: Studies in Child Welfare,* 1936, **12,** No. 3, pp. 69.
18. WHITEHEAD, T. N. Leadership within industrial organizations. *Harvard Business Review,* Winter 1936, pp. 161–171.
19. WHITEHEAD, T. N. *Leadership in a free society.* Cambridge, Mass.: Harvard University Press, 1936. Pp. xv+266.

[MS. received August 7, 1939]

MULTIPLE-VARIABLE DESIGN FOR EXPERIMENTS INVOLVING INTERACTION OF BEHAVIOR

BY RICHARD S. CRUTCHFIELD

Mount Holyoke College

AND

EDWARD C. TOLMAN

University of California

Traditional principles of experimental design require that but one independent variable be studied at a time, while all other conditions are held constant. This single-variable design insures unambiguity of results, since all variance in the data (other than 'chance' variance) is ascribable to the one source. Psychological research has depended almost exclusively upon this technique, and many textbooks of psychology go so far as to describe this single-variable design as the *only* proper one for fulfilling the prerequisites of rigorous experimental procedure.

It has been fully recognized by Fisher (**2**, **3**) and others, however, that in many cases such single-variable design may quite properly be superseded by a more profitable *multiple-variable* design. Multiple-variable design provides for the simultaneous study of two or more variables. That is to say, as one independent variable is systematically varied, so at the same time one or more additional independent variables are also systematically varied. Of course, all conditions other than those few under study are held constant.

Detailed description of the methods of multiple-variable design is given by Fisher (**2**). An illustration of its use in psychological research is given in an experiment by Crutchfield (**1**), in which five factors were simultaneously investigated with respect to their influence upon the expenditure of work in string-pulling by the rat.

The superiority of multiple-variable over single-variable

design is threefold. First, there is greater economy, in that several factors can be studied simultaneously, and with an equivalent degree of precision, but without any greater number of subjects than if one factor were studied alone. Second, there is a broader inductive basis for the generalization of results, since the effects of the various factors are evaluated over a wider range of conditions. Third, there is greater comprehensiveness, in that provision is made for the analysis of *interaction* of two or more of the factors.

By 'interaction' is meant a change in the effect of one factor associated with variation of another factor. Suppose, for instance, that a factor X is found to bear a certain functional relationship to a resultant behavior B, when another factor Y is held constant at value Y_1. Now if factor Y is varied through other values, Y_2, Y_3, etc., the relationship between X and B may change. This is called interaction of X and Y.

In this paper we wish to indicate the unique significance of multiple-variable design in the study of those areas of behavior where it is known or suspected that complex interaction of variables exists.

Tolman's theoretical system (**6**) takes cognizance of such complex interaction. According to his system, direct study of the isolated relationships between each of the independent variables or conditions, on the one hand, and the dependent variable—resultant behavior—on the other, is considered not feasible. Instead, certain mediate conceptual constructs (intervening variables) are developed, and these bridge the operational gap between the independent variables and the dependent variable. The underlying presupposition of this system is that the combination of variables is *not* simple and direct in nature, but is a complex synthesis of field-relations. It is here that Lewin's dynamic principles of topological and vector psychology (**4**, **5**) particularly pertain. And it is for the study and prediction of these complex interactions that Tolman's schematic 'sow-bug' (**7**) has been conceived.

We propose that the multiple-variable design is the most

fruitful experimental method for the evaluation of these interaction effects.

An illuminating illustration of how the multiple-variable design may be expected to function is provided in an experiment by Tolman now in progress. This study is concerned with vicarious-trial-and-error (VTE) at a choice point (7). Three variables are to be examined with respect to their effects upon the extent of VTE in a visual discrimination situation. These three factors are: (1) number of hours of food-deprivation; (2) width of gap to be jumped from choice-platform to either door; (3) degree of dissimilarity of the two stimuli to be compared.

Using conventional methods of design, each of these three factors would be studied separately, the other two being held simultaneously constant. Hence, a single curve of relationship would be established between values of each of the three factors and the related VTE scores. Each of these three curves would be valid only for the specific constant values of the other two influential variables. They could not be taken as valid for any other values of these variables, unless specifically demonstrated to be so. And this would necessitate a complete repetition of the experiment for each new set of values of these other variables.

By using multiple-variable design, however, the results would be made valid over a wide range of conditions, and without the necessity of increased replication of subjects.

Let us suppose that we are to study the three variables through a range of three values for each, that is to say, three degrees of food-deprivation, three widths of gap, and three degrees of dissimilarity of the two stimuli. Now according to the single-variable design, we would make three separate experiments, one for each variable. In each of these experiments there would be three sub-groups to be compared, corresponding, respectively, to the three treatments of that variable. If 12 rats were used in each sub-group, that would require 36 rats per experiment, or 108 rats altogether for the experiments. From the resulting data we would establish three simple independent curves. For example, with a given

width of gap, and a given degree of dissimilarity of the two stimuli, it might be found that VTE increases linearly with an increase in hours of food-deprivation.

In using multiple-variable design, however, the same number of rats, 108, could be tested by setting up 27 sub-groups of 4 rats each. These 27 sub-groups would correspond to the 27 possible combinations of treatment of the three factors, *i.e.*, three to the third power. It is evident that what would be obtained by this multiple-variable technique would be a *family* of curves relating each factor to the VTE score, each of these curves corresponding to a given set of values for the other two factors.

This experiment by Tolman is illustrative of a wide variety of problems in psychology to which the multiple-variable design is eminently applicable. There are numerous other examples. For instance, the extensive investigation by Warden and his co-workers (**8**) of the strength of animal drives might have benefited substantially by a use of multiple-variable design. Most of the relationships established in that study were simple curves relating the intensities of given drives to the frequencies of crossing an obstruction to the goal. All conditions other than the intensity of a particular drive were held constant. Hence, the generality of the curves which were obtained is extremely limited.

Had multiple-variable design been employed, other conditions—including the intensities of other drives besides the one being studied—could have been simultaneously varied. The relationships between intensity of drive and frequency of crossing would have been established, then, not simply for *one*, but for *several* sets of conditions. And of even greater significance for this whole study, direct information concerning the interrelationships or interaction of drives would have been available.

Multiple-variable design is appropriate, in general, for any comprehensive psychological study which is concerned with the generalized effects of a multiplicity of variables and their complex interactions.

REFERENCES

1. CRUTCHFIELD, R. S. The determiners of energy expenditure in string-pulling by the rat. *J. Psychol.*, 1939, 7, 163–178.
2. FISHER, R. A. *Design of experiments.* Edinburgh: Oliver & Boyd, 1937.
3. ——. *Statistical methods for research workers.* Edinburgh: Oliver & Boyd, 1936.
4. LEWIN, K. *Principles of topological psychology.* New York: McGraw-Hill, 1936.
5. ——. The conceptual representation and the measurement of psychological forces. *Contrib. Psychol. Theory*, 1938, Vol. 1, No. 4.
6. TOLMAN, E. C. The determiners of behavior at a choice point. PSYCHOL. REV., 1938, 45, 1–41.
7. ——. Prediction of vicarious trial and error by means of the schematic sowbug. PSYCHOL. REV., 1939, 46, 318–336.
8. WARDEN, C. J. *Animal motivation: experimental studies on the albino rat.* New York: Columbia Univ. Press, 1931.

[MS. received September 21, 1939]

THE BEHAVIOR MECHANISMS CONCERNED WITH PROBLEM SOLVING[1]

BY NORMAN R. F. MAIER

University of Michigan

Introduction

Because various problems may be solved by the functioning of quite different psychological processes, to speak of a problem-solving ability tends to confuse the real issue. As long as qualitatively different functions are involved in a situation, quantitative measurements are misleading and theoretically unsound. Quantitative measurement becomes significant only after single processes have been isolated.

Since problem solving designates an achievement we may speak of problem situations. The problem situation is characterized by the fact that it blocks behavior. If a situation is presented to an animal and elicits neither native nor acquired responses which remove the animal from the situation and if in addition the motivation is such as to demand such removal, then the animal is confronted with a problem. Problem situations are thus frustrating in the sense that a motivating condition cannot be resolved by an immediately available response.

Situations having the above qualifications may bring a variety of processes into play. It is my purpose to present an analysis of the processes which may be called out or brought to function in problem situations.

Variability in Behavior

A process which readily comes into play in a problem situation is that of variability in behavior. After investigation

[1] Communication from the animal laboratory supported by a grant from the Horace H. Rackham School of Graduate Studies of the University of Michigan. This paper supplements a previous paper (22) on the analysis of the learning function and was presented at the conference on the psychology of learning held at the University of Colorado, August 10–11, 1939.

of the behavior of various animal forms in a multiple choice problem, Hamilton (**6**, p. 43) concluded that "when a mammalian is confronted by a series of situations for which he is unable to discover and stereotype a specifically adequate and invariably successful mode of response he tends to vary his response in a manner which is less a species than an individual characteristic." A problem situation may be characterized by the fact that no specific stimulus is given which calls out particular native or acquired modes of behavior and as a consequence the process of variability clearly comes to expression and dominates over other processes. Krechevsky (**10**, **11**, **12**) has demonstrated the manner in which this process manifests itself and has shown by the comparison of normal and operated animals that the expression of variability is a function of the brain as well as of the situation. That this process is not random but in many cases can be scientifically analyzed was also demonstrated by Krechevsky (**8**) in his studies on hypothesis behavior. Although Krechevsky regards the hypothesis behavior as an indication of something akin to ideation I am inclined to disagree and use hypothesis as a term which describes the nature of the variability mechanism.

A problem situation, be it solvable or insolvable, thus gives rise to changing modes of behavior. Sometimes the changes are rapid and it is impossible to determine what aspect of the situation has dominated the animal's behavior; in other situations the animal frequently repeats a mode of response and in such cases one is able to state that, *e.g.*, the response is one based upon position or brightness. When one can discover the aspect of the situation which is determining the response Krechevsky speaks of hypothesis behavior and when this cannot be determined he speaks of variability. However, I see no reason for distinguishing between the two kinds of behavior except to recognize the difference in the frequency of changes in reaction.

For example, by the discrimination box method rats show hypothesis behavior when required to develop a preference for one member of a pair of cards. When the jumping technique is used, however, hypothesis behavior as a rule is not very

apparent. Since the evidence for hypothesis behavior depends upon repeated responses to the same aspect of the situation, rapid changes in response tendencies will not permit analysis in terms of hypotheses. Because mastery of the same task requires relatively few traits to learn if the jumping method is substituted for the Yerkes discrimination box, it is reasonable to suppose that the learning in the former case has been sufficiently speeded up to decrease repeated responses to the same aspect of the situation so that it is impossible to demonstrate clear-cut hypotheses.

We analyzed the hypothesis behavior of 8 normal and 8 partially decorticated rats while solving a discrimination problem and found that the normal animals show an average of 1.6 hypotheses whereas the operated animals show an average of 3. The trials to learn for these two groups are 102.5 and 127.5, respectively. Since Krechevsky found normal rats to have more hypotheses than operated animals our results indicate that hypothesis behavior was less apparent in our experiment because variability occurred too rapidly and not that hypothesis behavior failed to occur. Only in the operated animals was learning sufficiently slow to give clear evidence of several hypotheses. The normal rats varied their behavior so much that only now and then clear cut hypothesis behavior could be demonstrated.

It seems reasonable to suppose, therefore, that hypothesis behavior is a slow motion picture of variability and for this reason becomes a method for a more detailed study of variability.

Variability in behavior is thus the expression of a particular mechanism. The fact of variability may aid in the accomplishment of certain ends (*e.g.*, the solution of a problem), but such accomplishments are not descriptions of a mechanism since the variability may be present and yet not achieve the end, and the end may be achieved without variability having been expressed.

The extent of variability is changed whenever associations are formed. The persistence of unadaptive modes of behavior signifies that a certain reaction tendency is dominant.

Any mode of behavior continues to dominate unless some other reaction tendency supersedes. The association of reward and punishment with a response influences the relative dominance of the response, but the animal's native and acquired behavior repertoire as well as its native preferences must also be considered. Less intelligent animals change less rapidly than more intelligent ones because each additional experience changes their reaction tendencies less than those of superior animals.

From the above discussion it follows that the extent of variability and of hypothesis behavior should decline with decreased intelligence. Krechevsky (**9**, **10**, **11**, **12**) has shown in separate studies that brain injury reduces the extent of each.

Although variability in behavior is influenced by the formation of associations it is not dependent upon them since variability will continue so long as preferences are not acquired or the process making for variability is not replaced in some other manner. For this reason rats will alter their routes to food if this can be done without loss of reward and without too much additional travel (see Maier, **13** and Krechevsky, **22**).

Emotional excitement seems to be a physiological condition which inhibits the process making for variability. Hamilton (**6**, p. 31) stated that perseverance reactions tend to appear in emotional situations and points out that an adult "trapped and badly frightened in a burning hotel, will rush madly again and again to a part of the building which obviously will not afford escape." Recent experiments on frustration bring further evidence to bear on this point. A frustrated animal in many cases resorts to an earlier habit or natural tendency (see Hamilton & Krechevsky, **7**; Muenzinger & Wood, **26**; Everall, **5**; Sanders, **28**) and repeats this behavior, thus failing to achieve an adjustment in the new situation. Repeated frustration actually fixates modes of behavior more firmly than does continued reward for the same number of trials (Maier and Glaser, in preparation).

A problem situation therefore must not arouse emotional reactions if variability is to be effective. Even continued

failure in a situation may eventually inhibit variability and we may say that the animal has given up. Inactivity in such cases replaces the variability in behavior.

The psychological mechanism underlying variability often leads to the solution of a problem. When adjustment becomes learned the situation ceases to be a problem and the variability in behavior is replaced by a specific pattern.

The above mechanism accounts for solutions which seem to be achieved by accident and by trial and error.

Equivalence Reactions

A second kind of problem solving depends upon the mechanism which makes for equivalence reactions. Suppose situation ABC through previous experience gives rise to response X. The animal is now confronted with situation MBO. If response X occurs we say the two situations are equivalent but they are not equal because the animal can differentiate the two situations and expresses this differentiation by showing much more hesitancy in the second than in the first situation. Problem solving of this sort is usually treated under the term transfer of training. Perception is the essential process in such transfer reactions. Only when the B in the first situation is like the B in the second can we expect the transfer to appear. Individuals will differ in their equivalence reactions depending upon what aspect of the first situation determined the nature of the response.

For example, when rats were trained to discriminate a large black circle from a small black circle both on gray backgrounds, Maier (**21**) found that some learned to react to a brightness difference while others learned to react to a size difference. When confronted with a new set of cards in which brightness and size were conflicting, some of the animals expressed one preference and some an opposite preference. Both showed comparison behavior and hesitancy, but since the new situation contained features which aroused old habits, a specific—though different—response eventually appeared in each case.

The above situation is a problem situation in the sense

that no response is immediately available and none becomes available until a perceptual organization appears which arouses an old mode of behavior. In comparison with the first type of problem solving we may here speak of variability in perception. Experiments on the transfer of training primarily utilize this mechanism.

This type of process also makes possible what has been called concept formation. Training is merely continued until a limited aspect of a situation calls out a certain response. If this aspect happens to be the three sides of a geometrical figure, then we say the animal has formed the concept of triangularity. The animal has not generalized; rather it has formed a reaction to a limited aspect of the training situation. For a more detailed analysis see Maier and Schneirla (**24**).

Spontaneous Integrations

The third manner in which a problem may be solved is by the integration of two or more separate past experiences. This process is different from that of association in which past experiences become integrated because of contiguity. The first experience of contiguity between the elements to be associated is dependent upon the physical properties of the situation and association formation therefore is dependent upon a previous experience of this sort. Thus in a maze a blind alley may become associated with a particular junction because one follows the other. In other cases variability may serve to produce the first experience of contiguity. In contrast to such learned combinations spontaneous integrations also appear and these are produced by an altogether different process. It is this process that I have designated as reasoning.

I believe that it is basically different from learning for the following reasons:

1. The processes of variability and association may be excluded in experimental situations, and solutions occur nevertheless. This requires either a revision of the concepts of association and variability or the postulation of an additional process.

2. Since it has been shown (Maier, **16**) that the same brain injuries which reduce the ability to integrate experiences spontaneously may leave the association mechanism intact, acceptance of both types of mechanisms is logically sound.

3. The lack of correlation between scores on typical learning problems and reasoning problems indicates that the processes are independent variables (Campbell, **2**).

4. Young rats are superior to adults in learning problems and inferior in reasoning problems (Stone, **29**, **30**; Maier, **17**, **19**).

5. Children under 6 years of age are inferior to adult rats in reasoning problems which fact indicates that the process involved matures relatively late (Maier, **18**).

6. In learning problems associations are built up in backward order (Maier and Schneirla, **24**); in reasoning problems they are built up in the forward order (Maier, **19**).

7. A reasoning problem is solved as a whole and is not a matter of building up the solution in stages. Comparison of the VTE behavior (see Tolman, **31**; Muenzinger, **25**) of rats which could and could not solve a problem substantiate this point (Maier, **19**).

The above seven differences are experimental facts which must be recognized and explained. They do not depend upon a special manner of defining reasoning and learning since no single definition of a psychological process can logically account for opposite end results. To broaden the term learning to mean everything that is not native is to destroy what scientific meaning the term has achieved and make meaningless all of the psychological principles of learning.

In order to show how the integration of two separate experiences may give rise to something new and at the same time to satisfy the criteria of creative thinking it is desirable to tax the reader's inventive ability with the following problem.

It is well known that right and left train wheels and the axle between them are made of one solid piece of iron. This being the case, how do trains turn corners? In rounding a curve the outside wheel must travel farther than the inside

wheel and as a consequence (1) the outside wheel must make more revolutions; (2) the inside wheel must slip; or (3) the outside wheel must be the larger. The first possibility is excluded because the wheels are securely fastened; the second is contrary to fact; and the third is not reasonable since in such a case the train could not travel forward.

To construct a pair of such wheels which will turn a corner thus requires some inventive ability.

This was accomplished by constructing a wheel surface which has a larger inside circumference than its outside circumference. On a curve the centrifugal force throws the train to the outside causing the outer wheel to ride the track on its large diameter and the inside wheel to ride the track on its small diameter. On the straight track the centrifugal force is not brought into play and the wheels ride on their middle portions.

The experiences necessary for this invention consist of (1) large wheels can go farther than small wheels in the same number of revolutions; and (2) in going around a curve, a force tends to drive the vehicle to the outside. Both of these things are common knowledge, yet put together in the proper manner they constitute an important discovery.

If such problems are solved by an integration of familiar experiences, we must raise a question concerning the causes of such specific integrations. Two possibilities suggest themselves.

1. The experiences have in themselves the property of attracting and integrating each other (similarity, for example).

2. Forces outside the experiences bring about specific integrations.

I wish here to make a case for the latter possibility. In studies of problem solving in humans (Maier, 14) a process called 'direction' was postulated because it appeared that in addition to experience forces were operating in the individual which determined the kind of combinations he would make.

For example, in a problem requiring the construction of a hatrack in the middle of a room, when the materials consisted

of two poles and a table clamp, two distinct kinds of attempts were made. One consisted of building a hatrack with legs, the other of wedging the poles between the floor and the ceiling. Both approaches led to failure, but a suggestion which failed to be effective for the first type of approach was 100 per cent effective for the second. Other instances of this sort led me to postulate a dynamic condition in the individual which determined what integrations would or would not be made. Two types of dynamic conditions were in evidence and I have called these *habitual* and *new* directions.

Habitual directions are states in which the individual reproduces old solutions. With the process we can account for controlled or selective reproduction but not creative thinking. It is this type of problem solving which I believe corresponds to Krechevsky's (**8**) *hypothesis*, Ach's (**1**) *determining tendency*, Woodworth and Sells' (**33**) *atmosphere effect*, and in general what is called the idea in mind. Claparède (**3**) and Duncker (**4**) would agree that vague states in the individual determine the nature of the solution but would recognize that controlled reproduction is inadequate to explain the creative process.

New directions are less specific in that they do not result in reproductive solutions but give rise to combinations of old experiences. The directions are new in that they give rise to a new combination and hence a new product. In the human being these new combinations are accompanied by changes in meaning (Wertheimer, **32**) which I believe constitute the experience of insight. The new meaning arises from the fact that the old elements have a change in function in the new combination, and because the combination has never before been experienced the function is novel and thus introduces the surprise element. Insight thus becomes a consequence rather than a cause in problem solving.

The validity of this distinction between habitual and new directions is supported by some experiments now in progress. Miss Barbara Sherburne is testing the problem solving of two groups of college students differing in ability. By selecting students with intelligence test scores below the

10th percentile of college freshmen and above the 90th percentile she obtained two groups having marked differences in ability. These are being tested on the string problem (Maier, **15**) which requires the individual to tie the ends of two strings together each of which has one end fastened to the ceiling and the other end just reaching to the floor. When the subject takes hold of one string he cannot reach the other. To tie them together presents the problem of obtaining the string which is out of reach. The solutions fall into four patterns as follows:

1. Extending the reach by the use of a pole or some other object.
2. Increasing the length of one of the strings by fastening an extension to it.
3. Anchoring one string at a mid-point between the strings.
4. Fastening a weight to one and thus constructing a pendulum.

The first three of these I (**15**) have regarded as solutions which depend upon habitual directions. All of these things have been done before and involve no change in meaning. In each case the subject gets the string. In the fourth case the string must come to the individual. This solution arises when a pair of pliers and the string (which drags on the floor) combine to form a pendulum. In this case there is an element of surprise and a change in meaning since the tool changes to a weight and the string, which was too short, suddenly becomes too long and must be shortened. It might be claimed that a pendulum is an old experience. However, the pendulum does not exist until the two experiences are combined. We have given subjects work with a pendulum before introducing the problem and it has been an ineffective suggestion. Pendulums have not previously been used in this manner and are useless. But when the problem is solved the new combination is recognized as a pendulum.

That the fourth type of solution is different is shown by the fact that superior and inferior individuals use the first three types of solutions with equal frequency. The

groups differ markedly, however, in the frequency with which the pendulum solution is found.

It therefore appears that habitual and new 'directions' constitute two kinds of general functions which influence memory in problem solving. The *habitual direction* exerts a selective influence upon recall and concerning it there is considerable agreement since other psychologists have recognized the necessity of postulating a selective process. The function of the *new direction* is, however, less accepted and is often confused with the *habitual direction*. It is therefore necessary to bring additional evidence to support this mechanism. For this evidence we will again return to experimentation with rats.

In a recent investigation (Maier, **20**) the solving of a problem which required the integration of four separate experiences was analyzed. One of the problems was to determine whether the resulting integration or solution was brought about by the nature of the experiences and the situation, or whether an additional mechanism which produced the integration would have to be postulated.

For example, if experiences AX, B, CY, and D are integrated to form the pattern ABCD, are there forces within these experiences which draw them together (such as similarity) or must there be a background of forces established in the animal which acts upon the experiences and thus integrates them?

If the latter postulation is valid it follows that this background of forces must have properties different from the experiences themselves. It was believed that brain injuries which are known to have an effect upon memory might possibly have a different effect upon the integrative mechanism and in this way the memories and the forces which act upon them might be isolated.

The following experiment was therefore performed on two groups of rats: Group I was tested on the problem and given 20 tests. It was then subjected to cortical injury and on recovery was again tested. Group II was operated upon and tested without the pre-operative test.

Group I made a score of 55 per cent (chance expectancy being 25 per cent) before operation and 51.2 per cent correct after operation. The post-operative score of Group II was 29.9 per cent. Thus the post-operative scores of the two operated groups were significantly different, Group I having a score almost equal to its pre-operative score and Group II making a score not significantly better than chance.

The lesions of the two groups were found to be alike in size, locus, depth, and shape so that the difference in score was not dependent upon such features. The pre-operative training was therefore the only factor in the experiment which could account for this difference in score.

Pre-operative training may be regarded as having the following possible influences.

1. Score improves with learning and hence the loss through operation is offset by the practice effect. This possibility is, however, contrary to fact. The scores do not improve with practice, and any elements which are learned become a handicap. The best scores are actually made the first time the test is given.

2. Some influence other than memory carries through the operation. That this influence is not of the nature of a memory is evidenced by the fact that it survives a brain operation but cannot be acquired after one. Memories, on the other hand, are lost through operation but can be acquired after one. Thus the influence must be one not of the nature of memory traces, but rather one that depends upon some additional mechanism which is capable of acting upon memory traces.

It is my belief that what we have called *direction* is the expression of such a mechanism. It appears that a specific *direction* is established in the animal in a problem situation and is the force which integrates memories in a particular manner without being a memory itself. This condition is dependent upon adequate cortical tissue for its establishment but once the force is set up it can survive an operation.

That such a process exists is supported by the fact that

rats capable of solving the problem behave differently at the choice points than do animals incapable of solving the problem (Maier, **19**). Apparently the solution occurs before the animal actually begins to run in the situation. A successful animal solves the problem as a whole rather than by solving parts of it as it progresses.

Our theory of *direction* thus postulates that memories can be manipulated by a background of forces analogous to the manner in which the perception of a figure is dependent upon properties of the background as well as the properties of the figure. This means that memories are subject to modification not only by other memories as has been demonstrated by Wulf (**34**) and Restorff (**27**) but also by other forces within the animal which in turn depend upon the problem situation and the individual. Further experimental tests of the manipulation of memories in problem situations must be performed to substantiate further this theory.

Conclusion

In order to account for the various ways in which problems are solved it has been necessary to postulate at least three psychological processes. The terms variability, equivalence reactions, and reasoning seem adequate terms for designating these processes. I have attempted to describe these processes and to show that they can be independently measured and studied. To designate the process does not constitute an explanation, but the delimitation of processes is an essential first step in arriving at one.

Because most problem situations, *e.g.*, problems in logic and mathematics, may involve any combination of these three processes it is necessary so to design problems that these processes may be controlled. It is likely that the inadequate design of problems has been responsible for so much confusion in this field of psychological investigation. Each experimenter chose his own pet problem and then generalized on the nature of problem solving. Thus the use of certain problems led to the notion that problem solving was a matter of trial and error and could be reduced to

learning; transfer of training experiments supported the notion of generalization and concept formation; and detour experiments supported the insight hypothesis. Each type of experiment was supposed to investigate the same process and as a consequence the data seem to be contradictory. If we recognize that problem situations may be solved by quite different processes the basic contradictions disappear, but new difficulties arise. The experimenter then has more variables to control and he must use ingenuity in designing his experiment. Because problem solving studies have in general lacked the aspect of controlled experimentation it has frequently but justly been cited as a barren field of inquiry. By recognizing the fact that problem solving ability is the resultant of a number of processes the quality of experiments should greatly improve.

It might also be added that the use of the introspective technique can never be a substitute for controlled experimentation. It has been shown (Maier, **15**) that the crucial act of integration is lost to the reasoner, consequently the report of the experience of solving a problem is most incomplete. Furthermore the introspective technique allows the data to become loaded with rationalizations which are always present when descriptive accounts of vague experiences are requested. Knowing this, the social psychologist does not study the development of attitudes by asking subjects how they achieved the attitude. The subject can justify his attitudes and the manner in which he solves a problem but for the psychologist to accept such verbal reports, no matter how objectively they may be dressed, seems scientifically unsound, except in cases where first approximations and leads may be suggested. Experimenters utilizing introspective technique also find their theories involved and verbal in nature and have difficulty in relating their findings to subhuman data.

BIBLIOGRAPHY

1. Ach, N. *Über die Willenstätigkeit und das Denken.* Göttingen: Banderoeck, 1925. Pp. 294.
2. Campbell, A. A. Community of function in the performance of rats in alley mazes and the Maier reasoning apparatus. *J. comp. Psychol.*, 1935, **19**, 69–76.

3. Claparède, Ed. La genèse de l'hypothèse, Etude expérimentale. *Arch. de Psychol.*, 1934, **24**, pp. 153.
4. Duncker, K. *Zur Psychologie des produktiven Denkens.* Berlin: Julius Springer, 1935. Pp. 135.
5. Everall, E. E. Perseveration in the rat. *J. comp Psychol.*, 1935, **19**, 343–368.
6. Hamilton, G. V. A study of perseverance reactions in primates and rodents. *Beh. Monogr.*, 1916, **3**, pp. 65.
7. Hamilton, J. A., & Krechevsky, I. Studies in the effect of shock upon behavior plasticity in the rat. *J. comp. Psychol.*, 1933, **16**, 237–253.
8. Krechevsky, I. 'Hypothesis' versus 'chance' in the pre-solution period in sensory discrimination learning. *Univ. Calif. Publ. Psychol.*, 1932, **6**, 27–44.
9. ——. Brain mechanisms and 'hypotheses.' *J. comp. Psychol.*, 1935, **19**, 425–468.
10. ——. Brain mechanisms and variability: I. Variability within a means-end-readiness. *J. comp. Psychol.*, 1937, **23**, 121–138.
11. ——. Brain mechanisms and variability: II. Variability when no learning is involved. *J. comp. Psychol.*, 1937, **23**, 139–163.
12. ——. Brain mechanisms and variability: III. Limitations of the effect of cortical injury upon variability. *J. comp. Psychol.*, 1937, **23**, 351–364.
13. Maier, N. R. F. Attention and inattention in rats. *J. genet. Psychol.*, 1930, **38**, 288–306.
14. ——. Reasoning in humans: I. On direction. *J. comp. Psychol.*, 1930, **10**, 115–143.
15. ——. Reasoning in humans: II. The solution of a problem and its appearance in consciousness. *J. comp. Psychol.*, 1931, **12**, 181–194.
16. ——. The effect of cerebral destruction on reasoning and learning in rats. *J. comp. Neurol.*, 1932, **54**, 45–75.
17. ——. Age and intelligence in rats. *J. comp. Psychol.*, 1932, **13**, 1–6.
18. ——. Reasoning in children. *J. comp. Psychol.*, 1936, **21**, 357–366.
19. ——. A further analysis of reasoning in rats. II. The integration of four separate experiences in problem solving. *Comp. Psychol. Monogr.*, 1938, **15**, 1–43.
20. ——. A further analysis of reasoning in rats. III. The influence of cortical injuries on the process of 'direction.' *Comp. Psychol. Monogr.*, 1938, **15**, 44–85.
21. ——. Qualitative differences in the learning of rats in a discrimination situation. *J. comp. Psychol.*, 1939, **27**, 289–331.
22. ——. The specific processes constituting the learning function. Psychol. Rev., 1939, **46**, 241–252.
23. Maier, N. R. F., & Glaser, N. M. Unpublished data.
24. Maier, N. R. F., & Schneirla, T. C. *Principles of animal psychology.* New York: McGraw-Hill, 1935. Pp. 529.
25. Muenzinger, K. F. Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency. *J. genet. Psychol.*, 1938, **53**, 75–86.
26. Muenzinger, K. F., & Wood, A. Motivation in learning. IV. The function of punishment as determined by its temporal relation to the act of choice in the visual discrimination habit. *J. comp. Psychol.*, 1935, **20**, 95–106.
27. Restorff, H. von. Über die Wirkung von Bereichsbildung im Spurenfeld. *Psychol. Forsch.*, 1933, **18**, 299–342.
28. Sanders, M. J. An experimental demonstration of regression in the rat. *J. exper. Psychol.*, 1937, **21**, 493–510.

29. Stone, C. P. The age factor in animal learning: I. Rats in the problem box and the maze. *Genet. Psychol. Monogr.*, 1929, **5**, 1–130.
30. ——. The age factor in animal learning: II. Rats on a multiple light discrimination box and a difficult maze. *Genet. Psychol. Monogr.*, 1929, **6**, 125–202.
31. Tolman, E. C. The determiners of behavior at a choice point. Psychol. Rev., 1938, **45**, 1–41.
32. Wertheimer, M. Über Schlussprocesse im produktiven Denken. *Drei Abh. z. Gestalttheorie*. Erlangen: Philosophique Akademie, 1925. Pp. 164–184.
33. Woodworth, R. S., & Sells, S. B. An atmosphere effect in formal syllogistic reasoning. *J. exper. Psychol.*, 1935, **18**, 451–460.
34. Wulf, F. Über die Veränderung von Vorstellungen. *Psychol. Forsch.*, 1922, **1**, 333–389.

[MS. received September 21, 1939]

THE USE OF THE TERM "LEVEL OF ASPIRATION"

BY JOHN W. GARDNER

Connecticut College

The term *Anspruchsniveau* was first introduced into the experimental literature about nine years ago in the work of Lewin and his students, but gained little currency in this country until the appearance in 1935 of Lewin's *Dynamic Theory of Personality*. The term was translated into English as 'level of aspiration,' and in the past several years has gained widespread use. It has appeared in textbooks and other systematic treatises, in discussions of social psychology, motivation and personality. Conversationally, psychologists are increasingly inclined to make use of it as a provocative new term adding spice to old discussions of success and failure, inferiority attitudes, and human motivation. Under the circumstances, a brief discussion of the use of the term 'level of aspiration' seems to be in order.

The concept of 'level of aspiration' was first formulated by Dembo (**4**, p. 50–56), one of Lewin's students, in the course of an experimental investigation of anger. Her experimental situations employed frustration as a means of evoking anger, the subject being required to perform some task which was either extremely difficult or completely impossible. A by-product of her investigation was the discovery that when a required goal is too difficult the subject will set up an intermediate goal which is easier than, but a step toward, the required goal. This intermediate goal she termed the subject's *momentary level of aspiration.*

It is not the purpose of this paper to review the experimental work on level of aspiration, but it will serve as a useful introduction to the present discussion to describe very briefly a few of the early findings.

Hoppe (**10**), who published the first major study of level

of aspiration, states that experiences of success and failure in the course of an activity will depend upon whether or not the subject attains these intermediate goals which he himself sets. If the subject attains the intermediate goal (momentary level of aspiration) which he has set himself, he experiences the performance as a success, regardless of whether or not he has achieved a perfect score or attained the goal set by the experimenter. On the other hand, if the subject's performance does not come up to his momentary level of aspiration, he experiences the performance as a failure. Hoppe found that the level of aspiration shifts frequently in the course of an activity. If the individual's performance comes up to his level of aspiration, the latter tends to shift upward; if he fails to attain his aspiration-level on a given trial, the aspiration-level is likely to be lowered.

Of most importance to the work of later American investigators was Hoppe's observation that marked individual differences in aspiration-level appear, even in the course of simple laboratory tasks, and his suggestions that these individual differences may be diagnostic of important personality differences. He found, for example, that an individual may habitually set intermediate goals (*i.e.*, levels of aspiration) which are far above his level of performance (thus courting failure-experiences) even though his rate of improvement offers no grounds whatever for such high expectations. Another subject may consistently demand so little of himself (*i.e.*, be so careful to set intermediate goals within easy reach) that he always provides himself with the presumably satisfying experience of doing "better than he expected." Hoppe suggests that such individual differences in level of aspiration reveal differences in ambition, prudence, courage, self-confidence, courage to face reality, etc.

With this brief introduction, we may proceed to a consideration of the topic at hand. What do we mean when we speak of an individual's level of aspiration? How may the term be justifiably used in systematic discussions of motivation or personality?

Hoppe's delineation of what he understands by the term *level of aspiration* is contained in the following passage:

> The subject . . . always undertakes the task with certain demands (Ansprüchen), which can change in the course of the activity. The totality of these constantly shifting, now indefinite, now precise, expectations, goal-settings or demands in connection with one's own future performance, we shall term the level of aspiration of the subject (10, p. 10).

The meaning of the term level of aspiration bears an intimate relationship to the methods used in determining level of aspiration. Hoppe's method of determining level of aspiration is based upon a crude inferential technique. A subject's aspiration-level may be inferred, he asserts, through the simultaneous use of the three following lines of evidence: (1) the spontaneous remarks of the subject, (2) the occurrence of success and failure experiences, and (3) the way in which the subject 'goes at' the task. The relevancy of the second line of evidence involves the assumption that success and failure experiences are dependent upon the attainment (or failure to attain) one's momentary level of aspiration. If this assumption is valid, then the occurrence of an experience of success would permit the inference that the subject's aspiration-level on that trial was at least no higher than he had actually performed; and the occurrence of feelings of failure would permit the inference that the aspiration-level lay somewhere above the subject's actual level of performance on that trial.

Later investigators, while deriving much benefit from Hoppe's important exploratory work, have found it necessary to reject his methods of determining aspiration-level on the grounds of lack of precision and objectivity. Jucknat (11) offers a more precise method. She made use of a series of ten paper-and-pencil mazes, arranged in order of difficulty. The more difficult a maze was the larger it was, so that when the mazes were laid out in orderly arrangement on a table it was easy for the subject to recognize at a glance that he was facing

a series arranged in order of difficulty. The subject was told to choose whichever maze he wished and to start working on it. In doing so he automatically revealed his level of aspiration. Thus if his first choice was Maze No. 2 (next to the easiest in the series), his aspiration-level score for the first trial was 2.

The idea of framing the situation so that the subject automatically reveals his level of aspiration without verbal commitments is an appealing one, and represents an unquestionable advance over Hoppe's method from the standpoint of objectivity. It is important to note, however, that in employing this method Jucknat still clings to the hope that she is dealing with the level of aspiration as envisaged by Hoppe. The point is worth considering because it typifies a recurring error in aspiration-level investigation, and because, from a broader viewpoint, it represents a widespread tendency in the study of personality. Students of personality are increasingly eager to employ precise, objective, quantitative methods, but are still notably reluctant to content themselves with the data which such methods yield. The level of aspiration as obtained in Jucknat's situations is not the level of aspiration as conceived by Hoppe. It is clear throughout Hoppe's work that when he uses the term level of aspiration he is referring to the individual's true inner aims and expectations. But as soon as an individual is in a position in which his aspirations are automatically revealed (as in Jucknat's situations), social factors alter the picture. It is a commonplace that an individual's verbal behavior is no valid indicator of his inner attitudes; but the principle will bear re-iteration in its extension to non-verbal behavior. Individuals commonly 'edit' those aspects of their behavior, verbal or non-verbal, which are open to public inspection. When secretly cherished aims are jacked to the surface and forced into behavioral expression, they too become subject to this editing.

These statements are not to be interpreted as a condemnation of Jucknat's method, but only as an argument

for explicit recognition of the sort of data which the method yields. In short, if Jucknat wishes to deal with data thus obtained under the term *level of aspiration*, then the latter should be re-defined as, let us say, "that level in a difficulty scale at which the subject is willing to test himself in the presence of the experimenter."

That such a re-definition is necessary under such circumstances was recognized by Frank (**5**) in his own attempt to provide a more precise method of dealing with levels of aspiration. After each trial in a given task, Frank told the subject his performance score and then asked the subject how well he intended to do next time. He then re-defines level of aspiration as follows:

> The term *level of aspiration* . . . is defined as *the level of future performance in a familiar task which an individual, knowing his level of past performance in that task, explicitly undertakes to reach* (**5**, p. 119).

In his earlier papers Frank makes no comment upon this sharp divergence from the original definition, but in his most recent discussion (**12**, p. 465) he states that the reformulation was undertaken "with a clear conscience in the conviction that the concept remains essentially the same as that used by Hoppe." In the light of the present discussion, such a statement cannot be accepted. Hoppe conceives the level of aspiration as the totality of certain highly subjective demands, aims and expectations. With Frank the level of aspiration is the goal which the individual *says* he is undertaking to reach. This amounts to considerably more than a difference in emphasis; the two investigators are talking about two different things. Equating the two involves the assumption that the subjective aims and expectations of which Hoppe speaks are given direct, undistorted verbal expression.

Gould employs Frank's method, but makes clear her conviction that there is "no one-to-one relationship between what we might call 'true aspiration-level' and the quantitative measures of aspiration-level" (**8**, p. 275). Following the experimental sessions, Gould conducted lengthy interviews

with her subjects, and obtained a wealth of interesting and important material. However, it is interesting to note that while she is keenly aware of the inadequacy of the explicit statements of aspiration-level made in the experimental situation as reflections of inner aspirations, she is rather confident in making generalizations concerning the 'inner level of goal-strivings' (*i.e.*, true aspiration-level) on the basis of the statements made during the interview. For example, after reviewing some of the data of the interview, she states that "explicit estimates yield more rapidly to changes than does the motivating level of goal-strivings. . . ." (**9**, p. 46). How the interview statements can provide the basis for such confident knowledge of the motivating level of goal-strivings (*i.e.*, true level of aspiration) is not entirely clear. It is conceivable, to be sure, that considerations of rapport make the interview material less artificial than the experimental material; but in the absence of external evidence to this effect one may question the consistency of permitting the verbal expressions of the interview to achieve the status of valid indicators of inner attitudes, when the verbal expressions of the experimental situation are denied this status.

The issue may now be summarized: Hoppe thought of the level of aspiration as the totality of certain highly subjective aims. Later investigators rejected his methods as inadequate for getting at these aims, and set up more precise and objective methods. These methods were assumed by Jucknat and Frank to be getting at the level of aspiration as originally defined, but actually they were measuring precisely as much as the individual was willing to make public concerning his aims and no more.

The question arises, then, as to the extent to which this constitutes a serious weakness in aspiration-level research. It would seem, at first glance, that we must either give up the research until more piercing methods are developed, or go ahead using the methods at hand with the toughminded but nonetheless wistful assertion that we are dealing with 'manifest' level of aspiration because we simply have no way of

getting at the 'true' level of aspiration. Actually, neither alternative states the true case. Such feelings of methodological frustration should be dissipated by consideration of the following question: Is there any 'true' level of aspiration? In the opinion of the present writer the early speculation concerning the nature of the phenomena involved in this line of experimentation produced an *ignis fatuus* which has misled most investigators since. What reason is there to believe, after all, that Hoppe's inner "totality of . . . expectations, goal-settings or demands" is so structured as to warrant the term *level* of aspiration? Might not the subject's striving frequently be characterized by nothing more than an amorphous desire 'to make a good performance' or 'to do better than last time' (which implies a *region* of aspiration, perhaps, or a *direction* of aspiration, but not a level whose height is susceptible to quantitative evaluation)? Might not an individual on a given trial occasionally free himself completely of the scoring scale and concern himself with the improvement of his stance, with the discovery of weaknesses in his technique, with attempts to relax, etc? Might not an individual in a task such as dart-throwing entertain at one and the same time a wild hope that he will make a perfect hit and a more prudent hope that he will at least hit the target, with perhaps an additional, self-conscious hope that he will not appear too awkward in the eyes of the experimenter? In other words, is there not considerable likelihood that the individual's aims on a given trial are manifold, fluctuant, ephemeral, and differing qualitatively as well as quantitatively, with those aims which involve a specific score often giving way to aims which cannot possibly be described in terms of score values? And may not any or all of these aims, qualitative or quantitative, specific or vague, operate more or less concurrently in a kaleidoscopic pattern, with one aim dominant one moment and another the next? How, then, out of this maelstrom of aims, some quantitatively describable and others not, and all of them contributing to a pattern which is certainly not at present quantitatively conceivable, can we formulate a neat concept of level of aspiration?

There are true aspirations, of course, but what reason have we to suppose that they can justly be described in terms of a *level?* The motivation of such a belief is more obvious than the evidence in its favor. 'Level' implies quantifiability. It means that aspiration can take its place in neat, dynamic formulae such as May and Doob's *Motivation is a function of the discrepancies between level of aspiration and level of achievement* (**12**). It is far from our intention to discourage attempts to deal with true, inner strivings; but the way to bring such complicated material within the realm of measurability is not to endow it first with quantifiable characteristics which are utterly spurious.

There is, then, one and only one meaning to the term *level of aspiration:* it can only refer to a quantitative indication which an individual makes concerning his future performance in an activity. Such quantitative indications concerning future performance occur sporadically in certain appropriately structured life situations (*e.g.*, on the golf course, in a shooting gallery) but their systematic evocation demands a specially designed experimental situation. This 'specially designed experimental situation' has two important features. In the first place, it demands that the subject make some public indication of what he aims to achieve. Presumably, the subject offers whatever variety of information on this topic it best suits his momentary purpose to give. It may be true or false. It may be true on one trial and false on the next. And after what we have said concerning the subject's true aims, it should be apparent that in many cases the subject's publicly indicated aims may not be appropriately described as either true or false (since there is no underlying truth to adhere to) but only as artificial.

The second important feature of the situation is that the subject is required to put this information concerning his aims in quantitative terms.

Out of these two requirements of the experimental situation—that the subject make some public indication of his aims, and that he make this in quantitative terms—we obtain

the concept of level of aspiration. It is manufactured in the laboratory.

Used in this very narrow sense, level of aspiration is a truly quantitative concept. Used in any other sense, it is not only pseudo-quantitative but mythical. It may be objected that in narrowing the concept down to the point where it means something, we have narrowed it down to the point where it means nothing. The only convincing answer to this objection is to be found in the recent experimental literature on the subject (**1**, **2**, **6**, **7**, **14**), which indicates that even in its strictest meaning the concept represents one of the most interesting personality variables that has been uncovered in recent years.

We have concerned ourselves with what Lewin would term the empirical coordinating definition of the concept of level of aspiration. We would agree with Lewin that knowledge of the 'conceptual properties of the construct' is equally desirable; but the establishment of these conceptual properties remains the task of future experimenters in the field, and will not be hastened by reading more into the concept than is already there.

REFERENCES

1. Anderson, H. H., & Brandt, H. F. Study of motivation involving self-announced goals of fifth grade children and the concept of level of aspiration. *J. Soc. Psychol.*, 1939, **10**, 209–232.
2. Challman, R. C. The effect of failure on level of aspiration (to be published).
3. Chapman, D. W., & Volkmann, J. A social determinant of the level of aspiration. *J. abnorm. soc. Psychol.*, 1939, **34**, 225–238.
4. Dembo, T. Der Ärger als dynamisches Problem. (Untersuchungen zur Handlungs- und Affektpsychologie. X. Ed. by Kurt Lewin.) *Psychol. Forsch.*, 1931, **15**, 1–144.
5. Frank, J. D. Individual differences in certain aspects of the level of aspiration. *Amer. J. Psychol.*, 1935, **47**, 119–129.
6. Gardner, J. W. Aspiration-level in response to a prearranged sequence of scores. *J. Exper. Psychol.*, 1939, **25**, 601–621.
7. ——. The relation of certain personality variables to level of aspiration. *Psychol. Bull.*, 1939, **36**, 540.
8. Gould, R. Factors underlying expressed "level of aspiration." *J. Psychol.*, 1938, **6**, 265–279.
9. ——. An experimental analysis of "level of aspiration." *Genet. Psychol. Monogr.*, 1939, **21**, 1–116.

10. Hoppe, F. Erfolg und Misserfolg. (Untersuchungen zur Handlungs- und Affektpsychologie. Ed. by Kurt Lewin.) *Psychol. Forsch.*, 1930, **14**, 1–62.
11. Jucknat, M. Leistung, Anspruchsniveau und Selbstbewusstsein. *Psychol. Forsch.*, 1937, **22**, 89–179.
12. May, M., & Doob, L. Competition and cooperation. Bull. 25. New York: Social Science Research Council, 1937.
13. Murray, H. A. *Explorations in Personality*. New York: Oxford Univ. Press. 1938. Pp. xiv + 761.
14. Sears, P. S. Levels of Aspiration in Academically Successful and Unsuccessful Children. Ph.D. Dissertation, 1939. On file at Yale University Library.

[MS. received July, 1939]

THING CONSTANCY AS MEASURED BY CORRELATION COEFFICIENTS

BY EGON BRUNSWIK

University of California

The term 'perceptual thing constancy' is applied wherever a certain type of perceptual response is found, under ordinary circumstances, to vary concomitantly with a certain kind of physical property of distant, or 'distal,'[1] environmental bodies (such as size, shape, reflectivity to light) rather than with the actual 'proximal' stimuli directly elicited by such distal properties, either on the retina or on some other receptor surface of the organism. The classical quantitative expression of co-variance between any pair of variables is the correlation coefficient. The present paper attempts to demonstrate, in a quite preliminary and non-technical fashion, the use of correlational analysis as a means of representation of the degree of perfection of perceptual thing constancy. As our example we have chosen, in the field of size-constancy, an experiment which was especially designed for the purpose.

I

A set-up of 15 cubes made of natural hardwood, ranging from 50 to 70 mm, was presented to the observer at frontal planes of 2, 4, 6, 8 and 10 m distance, three cubes at each distance. The lateral distance between neighboring cubes was approximately 80 cm. In the rear of the room, at a distance of 12 m from the observer, a comparison series of 13 cubes ranging from 30 to 90 mm with step-intervals of 5 mm was set up. All objects were placed on tables of usual height. The observer was seated on a slightly elevated chair and so had a complete view of the set-up. The sizes and distances of the cubes are schematically represented in Table 1.

[1] This term, recently adopted in this connection by Heider (8), is less open to misunderstanding than the term 'distant' which has too narrow spatial connotations.

The measured size of the cubes which is the distal stimulus and is labeled **b** (body-size) is indicated in bold-faced type. As can be seen, there are three cubes each of 50, 55, 60, 65 and 70 mm height, randomly distributed over the whole field.[2]

TABLE 1

EXPERIMENTAL ARRANGEMENT (Raw scores in italics)

Distance from observer (meters)		Data on the cubes (in mm)		
12		Thirteen comparison objects with sizes ranging from 30 to 90 mm		
	b	**70**	**55**	**50**
10	*e*	*82*	*60*	*55*
	p	84	66	60
	b	**55**	**60**	**70**
8	*e*	*63*	*69*	*82*
	p	83	90	105
	b	**65**	**50**	**65**
6	*e*	*80*	*61*	*77*
	p	130	100	130
	b	**60**	**70**	**55**
4	*e*	*73*	*84*	*66*
	p	180	210	165
	b	**50**	**60**	**65**
2	*e*	*59*	*74*	*80*
	p	300	360	390
		Position of observer O		

The figure in ordinary print, placed in the third row of each group of figures, indicates in each case the proximal stimulus value, that is to say, the actual retinal or 'projective' size *p* of the cube in question, in terms of the comparison series as related to the actual position of the observer. For example, the left rear cube, size 70 mm and at a distance of 10 m, occupies the same space on the retina of the observer

[2] Thus size constancy in this case might have been spuriously supported by 'central tendency' imposed by the comparison series. There is however sufficient evidence from other studies that this factor is not a decisive one.

as would an 84 mm cube at the distance of the comparison series, 12 m. Similarly, the right front cube, size 65 mm, is projectively equal to a 390 mm cube at 12 m distance; and so on. The values indicated are only approximate since lateral distortion has not been taken into account. This latter, however, from the point of view of our purpose, is of minor importance.

The observer was asked to take a natural, naive and unconstrained attitude, and to match each of the 15 cubes with the comparison series. Eight students were used as observers, each of them running through the experiment only once. The order of the judgments was systematically varied from observer to observer. The italicized figures in the middle of each group of figures give the perceptual (verbal) responses, namely the average estimates, *e*. For the purpose of our paper they may as well be considered as the result of one single experiment with one observer only, since at no stage of our considerations will the matter of individual or of time differences be considered nor will the correlation technique be applied to these latter.

In each of our fifteen instances, *e* is in an intermediate or 'compromise' position between *b* and *p*, usually much closer to the former than to the latter. This way of reacting is typical of most constancy experiments performed under ordinary conditions, yielding what has been labeled the phenomenon of 'approximate size constancy.' We may express the degree of this approximation by computing correlation coefficients between the three sets of variables.[3] The result is given in Table 2.

Two of these three coefficients contain the response-variable, *e*. *They indicate that estimated size shows a much higher degree of concomitant variation with the size of the distal environmental bodies than with the size of their projections upon the retina though these latter are essential in conveying body size to the organism.* The fact that r_{eb} is not unity, and r_{ep} not as

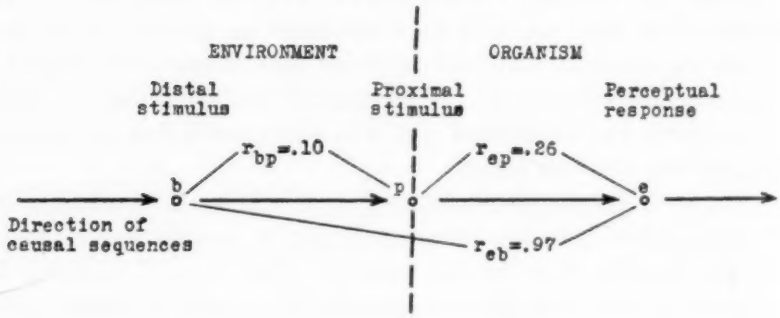
[3] Since our concern is only with demonstrating the principle, the numerical values given in Table 1 have been used in computing these correlations, instead of the probably more preferable logarithms of these values.

low as r_{bp}, is an expression of the lack of perfection of the constancy mechanism, which gives p, per se, an exaggerated emphasis.

The third coefficient, r_{bp}, has nothing to do with the observer's response but is concerned rather with the purely external interrelationship between the size of the distal body and the size of its retinal stimulus representation. This coefficient is very low. Indeed, when taken by itself, p is a very 'unreliable' cue for the size of the body 'causing' p (*cf.* **5**, **6**, **19**). Due to the fact, however, that large objects placed at various distances will on the whole elicit somewhat larger

TABLE 2

CORRELATIONS AMONG DISTAL STIMULI, PROXIMAL STIMULI, AND PERCEPTUAL RESPONSES



retinal images than small objects at similar distances, the correlation is not quite as low as would seem to be desirable from the standpoint of an experiment such as ours, *i.e.*, it is not quite zero. Since this tendency in our case (as often in nature) is insignificant in comparison to the distorting influence of the difference in distance, r_{bp} is low enough to be discarded in a first approximation. By slight changes in the arrangement it could, as *e.g.* for the purposes of a further demonstration, easily be brought down to zero.

The fact that the distal correlation *eb*, directly expressing far-reaching perceptual achievement, is high whereas the two correlations containing the mediating proximal link p are low may, when taken by itself, seem paradoxical. The

answer is, of course, that something has been left out of the picture, namely a group of proximal stimulus feature more or less correlated with the distance between object and observer which, duly brought together with p, will functionally explain the presence of a high correlation between e and b. The fact remains, however, that the proximal stimulus configuration mediating this correlation, or the constituents of this configuration, show a high degree of variability and flexibility, whereas e and b do not, and thus it is these latter which stand out as focal points of perceptual or organismic activity in general. It is these focal points, usually of vital importance, in the near or remote environment upon which a molar psychology should concentrate its efforts. A supplementary consideration of the general character and of the degree of the flexibility of mediation as expressed by low mediational correlations will help to complete the picture.

II

In recent years the degree of perfection of the constancy mechanism has repeatedly been represented by means of a constancy ratio, c, introduced by the author (**3**), whereby

$$c = \frac{e - p}{b - p}.$$

By the use of this formula, the degree of constancy can be computed for each of our cubes separately, yielding 15 c-values. They tend to be higher for the cubes near the observer than for those farther from him (for reasons which are of little bearing here). If computed from the logarithms of e, b and p instead of from the numerical values (*cf.* **17**, p. 344; **4**, p. 391), their average is .69.

The constancy ratio has several practical advantages. It is applicable, in principle, to any single pair of cubes set up at different distances and judged by an observer as equivalent with regard to apparent size. Furthermore, it expresses the constancy-achievement by one single index, c, relating e to both b and p at the same time. And it does so in such a way that the fixed values $c = 0$ and $c = 1$ always represent the

same kinds of outstanding achievement, namely complete lack and complete perfection of perceptual constancy, respectively.

The correlational technique suggested above possesses none of these advantages. It requires a whole set of bodies of various sizes and at various distances (not to mention requirements regarding the frequency distributions of these sizes and distances). These objects have either to be ranked among each other, or matched to a common scale, as was done in our experiment. The degree of achievement, or its contrary, deviation from the correct judgment in a certain direction, appears in distinctness only after a joint inspection of all three of the coefficients has been made, or only after some new index has been evolved. (In the simplest and crudest possible case, the latter could, for instance, be the difference between r_{eb} and r_{ep}, .71; which, by the way, in our particular case would be very similar to our average c-value, .69.)

On the other hand, as the author has pointed out previously (**4, 5**), the constancy ratio is too much bound up with certain rather incidental aspects of the experiment, such as the physical dimension in which the stimulus-situation is varied. It is of little use when the interplay of more than two variables is to be considered, and it does not isolate sharply the character even of these two variables (*cf.* **5, 1**). It is not limited to values between the ideal poles 0 and 1. For example, values above 1 also indicate deviations from perfect constancy as do those below 1 though in a different direction. This shortcoming has been especially emphasized by Koffka (**12**, pp. 227, 234).

In contrast, the use of correlation techniques, though clumsier at first sight, is an approach of much more basic significance. It brings the constancy problem within the scope of the standard instrument developed to disentangle complex causal textures of whatever nature. It helps in the giving up of the exclusive search for strictly univocal correspondences, in psychophysics as well as in stimulus-response psychology in general, by rendering more legitimate the carrying over from the psychology of individual differences to these

disciplines of what the author would like to call a 'deliberate lump-treatment' (7).

In doing so our approach goes beyond 'multidimensional psychophysics' of the merely proximal type, represented, for example, by the studies of Stevens (16, p. 70 *ff*; *cf.* also 20, p. 509), or Richardson (13). In this latter fairly recent type of research a perceptual response (such as the apparent loudness of a tone) has been studied as a function of more than one stimulus-variable (such as the frequency as well as the intensity of the underlying sound wave). Though the relationships found may justly be called complex and thus in a certain respect lacking univocality, yet the scope of such studies remains limited to the strictly lawful relationships which hold between a complex of proximal stimulus features and the response. Our aim, on the other hand, is to expand the investigation, regardless of the lack of a perfectly reliable representation of distal events in the proximal region, into the distal environment.

In so widening our scope, we are enabled to determine statistically the distal (along with the proximal) factors upon which the perceptual system has become fairly well focalized and which thus have become virtually the most effective determiners of the response. It is this type of far-reaching stimulus-response correlation which Holt (11, pp. 161 *f.*) and Hobhouse (9, p. 15) had in mind (*cf.* Heider, 8) and to which even earlier objective psychologists like Bechterev (2, p. 17) and Watson have occasionally referred, in an abstract way, and yet failed to deal with concretely, due to the difficulties caused by the lack of perfection inherent in such couplings.

This whole procedure should not be conceived as being limited to problems of perception, and to the use of correlation coefficients in the technical meaning of the term. Problems of overt behavior such as the obtaining of food by an animal are just as capable of quantitative correlational analysis of the relative importance of proximal and distal effects.

III

Our method differs from the most common use of the correlation coefficient in that the role of the individuals tested is

taken over by the objects in the physical environment of a certain individual organism, and the role of the tests applied to the individuals is taken over by various (in our case three) kinds of manifestations, or effects, of these objects, namely size as measured by applying a meter-stick directly to the object (*b*), size of the effect of the object on the retina of the organism in question (*p*), and size as it is perceptually anticipated by that organism (*e*).

Thus our method has little in common with the modification of procedure suggested by Stephenson (**15**), who interchanged individuals and tests, and whose method thus, in contrast to ours, remained in the realm of individual differences.

Likewise, our method differs from the application of the correlation technique and of factor analysis to any combination of distal perceptual tasks such as occurs in the studies of Thouless (**18**), Sheehan (**14**, pp. 52–56), and Hofstätter (**10**, pp. 27–33). There again the concern was with factors within the personality determining individual differences in perceptual achievement whereas the determining factors in our case are of an environmental nature.

As has been mentioned above, our method also transcends proximal multidimensional psychophysics, in that it includes the testing of distal determinants. It thus may become a tool for what might be called distal (and multidimensional) psychophysics.

In a certain way related to our approach are cases like that of Woodworth (**20**, p. 251) in which the emotions as expressed on a person's face have been correlated with judgments of these emotions given by a group of observers. The emotional state of another person is an environmental feature even a step farther distal than the size of a body. The chief difference when compared to our procedure lies in the fact that in the case just mentioned the distal stimulus only is considered, and the more proximal representations neglected. The same holds for cases in which the intelligence of a group of persons has been correlated with intelligence as judged from photographs of these persons.

Our procedure, on the other hand, deals with both the distal and the proximal, the achievemental and the mediational aspects. The outcome may lead to subordinating the latter to the former, but as far as the procedure is concerned both are taken into account and their relationship is examined.[4]

This article is no more than a sketchy beginning, with the purpose of demonstrating a general principle. All questions of detail and the further development of the statistical procedure such as the application of the analysis of variance, the isolation of the effective environmental factors, the fulfillment of the requirements for the application of the standardized correlational methods (such as normal distribution of the variables, etc.) and their possible modifications, the question of reliability of measures, etc., will have to be discussed in further publications of a more technical nature.[5]

Summary

An example has been given from the field of perceptual size constancy of the way in which correlational analysis may be helpful in determining quantitatively the degree to which the perceptual system—or the organism in general—is successful (under ordinary circumstances) in giving a specific response to, or in focalizing upon, a certain feature of the remote physical environment in spite of disturbances resulting from the incidental character of the proximal (retinal) stimuli mediating these distal environmental features. And it appears as characteristic of the constancy mechanism that high correlations of the response with distal factors may be accom-

[4] In a recent article (8), Heider has interpreted the author's conception of a 'psychology in terms of objects' as being limited to the distal aspects with exclusion of the proximal. Actually, this is not the case. Foci of correlations with organismic events will be recognized wherever they may be found, in the proximal as well as in the distal environment. Furthermore, even in the case of distal foci there is interest in the proximal aspects of mediation. This concern is best shown by the positive assertions made in this paper about the lowness of the correlations holding for a certain type of proximal events.

[5] The author is being aided in this work by a statistical discussion group under Professor Tryon. Some of the special problems will be worked out in collaboration with Mr. Robert Gottsdanker.

panied by low correlations with the mediating proximal cues. The author's ultimate aim is to establish a multidimensional psychophysics which will include the distal environment within its scope.

REFERENCES

1. Ansbacher, H. Perception of number as affected by the monetary value of the objects. *Arch. Psychol.*, 1937, **215**, 1–88.
2. Bechterew, W. *La psychologie objective.* Paris: Alcan, 1913, iii and 478 pp.
3. Brunswik, E. Zur Entwicklung der Albedowahrnehmung. *Z. f. Psychol.*, 1928, **109**, 40–115.
4. ——. Die Zugänglichkeit von Gegenständen für die Wahrnehmung und ihre quantitative Bestimmung. *Arch. ges. Psychol.*, 1933, **88**, 377–418.
5. ——. *Wahrnehmung und Gegenstandswelt.* Leipzig, 1934, xi and 244 pp.
6. ——. Psychology as a science of objective relations. *Phil. Sci.*, 1937, **4**, 227–260 (for errata, see same journal, 1938, **5**, 110).
7. ——. The conceptual focus of some psychological systems. *J. Unified Science (Erkenntnis)*, 1939, **8**, 36–49.
8. Heider, F. Environmental determinants in psychological theories. Psychol. Rev., 1939, **46**, 383–410.
9. Hobhouse, L. *Mind in evolution* (3rd ed.). London: Macmillan, 1926.
10. Hofstätter, P. R. Über die Schätzung von Gruppeneigenschaften. *Z. Psychol.*, 1939, **145**, 1–44.
11. Holt, E. B. *The Freudian wish.* New York: Henry Holt, 1915.
12. Koffka, K. *Principles of gestalt psychology.* New York: Harcourt Brace, 1935, xi and 720 pp.
13. Richardson, M. W. Multidimensional psychophysics. *Psychol. Bull.*, 1938, **35**, 659–660.
14. Sheehan, M. R. A study of individual consistency in phenomenal constancy. *Arch. Psychol.*, 1938, **222**, 1–95.
15. Stephenson, W. Correlating persons instead of tests. *Character and Personality*, 1935, **4**, 17–24.
16. Stevens, S. S., & Davis, H. *Hearing: its psychology and physiology.* New York: John Wiley, 1938.
17. Thouless, R. H. Phenomenal regression to the 'real' object. *Brit. J. Psychol.*, 1931, **21**, 339–359.
18. ——. Factor analysis in problems of perception. *Proc. Eleventh Internat. Congr. Psychol.*, Paris, 1937.
19. Tolman, E. C., & Brunswik, E. The organism and the causal texture of the environment. Psychol. Rev., 1935, **42**, 43–77.
20. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938, xi and 889 pages.

[MS. received August, 1939]

BASIC CONCEPTS AND PROCEDURES IN A STUDY OF BEHAVIOR DEVELOPMENT[1]

BY MYRTLE B. McGRAW

Columbia University

I. The Role of Concepts in Scientific Procedures

The time can still be recalled when students preparing for experimental work, especially in the biological sciences, were exhorted to 'stick to facts,' or 'let the facts speak for themselves.' Young investigators were made to feel that the business of science was much like that of exploration—the world of the unknown was there and it was almost the sole duty of the scientist to unearth these hidden facts. Once brought to light such facts would stand boldly evident as part of our body of scientific knowledge.

But those times have changed. Certainly the advancing guard of investigators today realize that concepts pertaining to the subject-matter and to the methods employed in manipulating or intellectualizing data are tools just as significant to experimentation and research as are the instruments of measurement or the actual facts obtained. Concepts of the investigator and those which have accumulated about the subject-matter are salient factors in the process of investigation, whether they are recognized as such or not, and they strongly influence the manner in which facts are obtained, the interpretation of the data, and the organization of knowledge about the data. It would therefore seem advisable to admit the role of personal and social concepts in scientific undertakings and to work toward their improvement and refinement as deliberately as one would set about improving a piece of apparatus. It should also be recognized that concepts and methods as tools of inquiry alter with time. Often these conceptual alterations are quite haphazard and uncul-

[1] From the Normal Child Development Study of the Department of Diseases of Children, Columbia University, and the Babies Hospital.

tivated; sometimes the usefulness of a particular concept is thereby utterly destroyed. A concept which was quite definite and precise in its origin may in the course of time acquire such a cargo of meanings that its original identity is lost. Others are vague and loose in their inception but attain specificity and definition as time goes on.

It is sometimes possible by arbitrary and deliberate demarcations to save a useful concept the unhappy lot of becoming a symbolic football in the game of scientific polemics. The concept of meter, for example, is one which has attained precision by virtue of national and international agreement. There is no specific entity called meter. The concept of 'meter' is often considered as a name for something concrete in existence, because it has existed for so long in our scientific thoughtways that its origin and actual import are frequently overlooked. Those investigators who are dealing with phenomena and relations for which no useful units of measurement, such as meter, are available would do well to pause and consider the manner in which such concepts of measurement in the phyiscal sciences gained their present degree of specificity. "The standard meter is a bar of platinum kept under as constant conditions of temperature and pressure as possible in the city of Paris. But if that were the whole story, the word *meter* would not have the connection with measuring it actually has. By itself the bar is just a particular bar and nothing else; it is neither a standard of measurement nor is itself measured. It is a measure of length because (1) all other rods of a meter's length in use anywhere in the world may be checked by being matched against it, and (2) because, and *only* because, these other rods are themselves used in matching still *other things*. It is just as true that the length of the bar of platinum (or any other measuring rod) is determined by its application in measuring cloth, walls, sides of fields, etc., as that the length of the latter is determined by comparison with it" (**2**). It should be borne in mind that the now specific concept 'meter' developed out of vague descriptive analysis of the relations between things in space and that it arrived at and retains its precise meaning by

virtue of international agreement. Furthermore, in its present use it represents not single units but a series of operations which form the basis for comparing relations between things in existence. This illustration is offered because a little retrospection on the origin and development of some of these precise units of measurement in the physical sciences would no doubt be of aid in bringing about greater precision in concepts of biological phenomena, especially in studies of behavior development where one is often dealing with a constantly changing phenomenon which does not lend itself readily to quantitative determinations.

II. Concept of 'Organism' as Representing Field of Inquiry

But before becoming too concerned over units of measurement or quantitative determinations, investigators in the field of behavior development should first of all consider the nature or quality of the phenomenon being investigated. For the moment may we say that any study of behavior development embraces an analysis of changes taking place within a living organism. At the mention of the word 'organism'—to say nothing of 'living'—we are beset with difficulties, and are immediately confronted with a concept in dire need of refinement. Recent literature is replete with discussion of the 'organism as a whole' and it is again and again emphasized that the functioning of the total organism is something quite different from the assemblying of parts or elements of which it is composed. But it can safely be said that the term 'organism' has not been subjected to a rigorous definition. Certainly the term 'organism' implies organization and organization is the essence of living matter. Bertalanffy emphasizes this point when he says:

> Since the fundamental character of the living thing is its organization, the customary investigation of the single parts and processes, even the most thorough physico-chemical analysis, cannot provide a complete explanation of the vital phenomenon. This (type of) investigation gives us no information about the co-ordination of the parts

> and processes in the complicated system of the living whole which constitutes the essential 'nature' of the organism and by which the reactions of the organism are distinguished from those of the test tube. But no reason has been brought forward for supposing that the organization of the parts and the mutual adjustments of the vital processes cannot be treated as scientific problems. Thus, the chief task of biology must be to discover the laws of biological systems to which the ingredient parts and processes are subordinate. *We regard this as the fundamental problem of modern biology* (1, p. 64–65).

Yet granting that an organism represents the organization of living systems, we still do not know, for experimental purposes, just what constitutes the boundaries or the scope of any particular organism. As the term is used, it seems to refer in most cases to the scope or field of inquiry rather than to any peculiarly identifying characteristic. It may be easy enough to identify a single cell floating around in plasma as an individual organism and one may undertake a study of the organization of nucleus and cytoplasm therein contained, but when one begins a study of the behavior of mammals, the business of identifying tangible boundaries of an individual organism is infinitely more complex.

Would it not be more practical, therefore, to accept an operational definition of the organism? That is to say, the phenomenon to be observed is 'organization,' and the limits, or the field of observation, are prescribed in accordance with each investigation. For some purposes this field, or scope of observation, may be a single cell, for others an individual as bounded by epidermis, or it may be a functional system as it operates within the epidermis, and for still other purposes it may be the universe. Actually there are no sharp lines which clearly demarcate an individual, or an organism, except as they are experimentally or operationally determined. Socially speaking, an individual represents merely nodal points of activity within a society and for some purposes the whole society may constitute the experimental organism. In much the same way that arbitrary limits can be prescribed around

nodal points within a society to signify a particular individual, so may arbitrary limits, for experimental and observational reasons, be set around nodal points of a psycho-motor activity. These nodal points of a psycho-motor expression may be called a behavior pattern, or an experimental organism, so long as it is the organization of neuro-muscular movements which concerns the investigator. Once the limits of observation have been deliberately established the changes in the organization of the activity over a period of time would comprise a study of behavior development.

Organization, then, is a quality varying in amount and common to all matter and energies. It may strain the imagination but certainly it is not altogether inconceivable that the concept of organization could be so defined that a unit of measurement could be applied which would indicate degrees of organization without regard to the particulars of the organism, similar to the way in which a mercury thermometer may be used to indicate changes in temperature of things which are qualitatively quite different. It was not until temperature was conceived as a quality permeating substances which were qualitatively unlike, that a device for obtaining symbolic measures of temperature in varying substances was feasible. Degrees, as indicated on a mercury thermometer, are not measures of temperature directly, but are merely symbols representing a quality which has been intellectually abstracted from the substance in which it exists (**2**). It is not necessary to extract the temperature of a particular substance in order to match it against the temperature of another. The parallel is obvious; it would not be necessary to dissect 'organization' from the organism in order to study its process of change, provided some proper symbolic indicator could be devised and applied as a measure of organization. Construction of such a symbolic system must await refinement in the concept of organization and its meaning with respect to living organisms.

In any study of behavior development a necessary procedure at the outset is the determination of the scope of inquiry. This field of inquiry is arbitrarily determined by

the investigator and the nature of the problem. In studying the development of reaching-prehensile behavior of the infant, for example, the object in the field of vision is just as much an integral part in the organization of the behavior as are the arms, fingers and eyes of the baby. And if one sets out to study the organization of this activity, the bottle for which the baby extends his arms and fingers is a part of the behavior function as well as the neuro-muscular movements which the baby makes. One manipulates arms and fingers quite differently when picking up a bowl of water from the way one does when trying to catch a fly. In that sense the object determines the configuration of neuro-muscular movements, and as such it might be considered 'an organizer' of behavior. But the manner in which a baby would pick up a bowl or catch a fly is still different from the way in which the adult would do it. It is these changes in manner of performance which reflect different degrees of organization in a given behavior activity. The boundaries of inquiry, that is to say, the experimental organism, having been deliberately prescribed, the next procedure involves an analysis of the changing degrees of organization within these boundaries.

III. The 'Longitudinal Method' of Collecting Developmental Data

Admission of the word 'development,' or changes in behavior, imposes upon the investigator the necessity of observing the same or comparable observational fields from time to time. It is the realization of this necessity which has evoked the current emphasis upon 'longitudinal' method in the biological sciences, especially in the field of child development. Along with this emphasis considerable confusion has accrued as to what is actually involved in a longitudinal method of investigation.

Dr. Palmer (**5**) in discussing this state of confusion with regard to the longitudinal method said:

> In a general way most workers in the field of growth and development would agree that it is possible to classify data dichotomously, one type being known as longitudinal and

another type being known as cross-sectional. Agreement may also be expected with respect to the principal characteristic which differentiates these two classes of data. Thus it is generally recognized that cross-sectional observations are those made on different individuals while longitudinal observations are those made on the same individuals. All of this, however, scarcely furnishes even working definitions and there seems to be considerable difference of opinion regarding more explicit definitions.

According to one concept which more or less expresses the views of some workers, the matter is handled by defining longitudinal data as follows: Longitudinal data consist of a long series of repeated measurements of the same characteristic observed on the same individuals. The broader aspects of this definition appear to be reasonably clear. For example, it is generally agreed that records of the weights of a group of children, each weighed on their birthdays for a period of years, constitute longitudinal growth data. It is clear in this case that the same characteristic (weight) must be measured several times. It is equally clear that successive weighings must be made of the same children.

The part of the above definition which causes difficulties concerns the idea that 'a long series of repeated observations' are required if data are to be called longitudinal. This idea leads, I believe, to difficulties as regards a clear concept of what longitudinal data are, as regards the usefulness of longitudinal data and also as regards methods of analysis of such data.

What is a 'long series' and what are 'repeated observations'? The implication is clear that some interval of time should or must elapse between repeated observations. It is implied also that many observations must be made. The question of how long the interval of time must or should be, or of how many repeated observations must or should be taken, is not explicitly stated so far as I know.

In an attempt to clarify these points, I should like to give and defend another definition of longitudinal data. This definition is as follows: Any datum consisting of two or more observations of the same characteristic observed on the same individual shall be considered longitudinal.

> In amplifying this statement, it may be emphasized that the interval between successive measurements may be any time whatsoever. In actual practice, the only limiting factor which is placed on the interval of time between measurements is that the interval have meaning in the particular problem under investigation.

Palmer's definition of longitudinal data is, in a sense, a refinement of our loose concept, but in another sense it is too broad to illuminate or spot the essence of a longitudinal study. According to his definition weight measured at birth and again when the child is ten years old would yield longitudinal data. Such data would not, however, reveal the way in which the weight increment occurred. Of course it is the manner in which data are handled and interpreted as well as the frequency with which they are collected that indicates their longitudinal merits. But it would be necessary to have at least three observations on any phenomenon in order to show its course of growth, unless we are going to admit that all development is in the order of a straight line.

It is in studies of growth and development that the longitudinal method gained eminence. Growth and development are concerned with change—change in form or configuration, or change in magnitude. Any truly longitudinal study must be conducted in such a way as to indicate the nature of the changes taking place in the observed phenomenon. It would seem, therefore, that a longitudinal study would be comprised of any series of observations on a changing phenomenon taken successively from the moment of inception until the changing characteristic attains stability or decline. The moment of inception and the state of stability or decline should be arbitrarily or operationally defined in terms applicable to the experimental procedures. It must, furthermore, be emphasized that the data should be taken in such a way that they can be handled intellectually and mathematically so as to take advantage of their longitudinal qualities, that is to say, handled in a way to show not only the fact of change but also the trend, velocity, and nature of such changes.

The crux of the difficulty in collecting longitudinal data

seems to hinge primarily upon the interval between observations. Palmer has suggested that the interval between measurements may be 'any time whatsoever' so long as the interval has meaning with respect to the problem being investigated. In most of the longitudinal studies published in the field of child development, the general assumption has been that the intervals between observations are adequate provided they embrace all changes of any detectable magnitude. This assumption certainly has logical meaning in so far as it bears directly upon the problem being investigated. On the other hand, if it is fully realized that it is the way in which the data are handled which determines the longitudinal merit of a study, and not the fact that data are collected in serial order, then it will be also recognized that the interval between observations is determined by the demands imposed upon the data for adequate mathematical and intellectual analysis. If one is dealing with a characteristic for which numerical values are obtainable and if the approximate rate of change in that particular characteristic is generally known, then the interval between measurements may be calculated at frequencies just sufficient to embrace all measurable changes. For example, it would be wasteful and laborious to measure the degree of bone calcification in the wrist of the same child every day. In the first place, the degree of change from day to day is incommensurable according to our present methods, and present methods of measurement do yield definite numerical values which can be manipulated for the purpose of indicating velocities, accelerations, regressions, and trends of development in that particular characteristic. It is wrong to make daily records on a characteristic of this type only because it is wasteful. On the other hand, if one is dealing with a changing characteristic for which no numerical or distinctly quantitative values obtain, or if one is dealing with a characteristic the general trend and developmental rate of which is not known, then the data must be accumulated at great frequency, even when no significant changes are taking place. The increased frequency of observations makes it possible to reduce the data to some symbolic form so that they

can be handled mathematically or intellectually in a way that would not be possible with the original facts. Certain types of data may be recorded longitudinally at sufficient intervals to include all significant changes in the developing characteristic, and yet the data will remain utterly useless for symbolic or intellectual manipulation. For example, it is possible to record on a particular baby that at the end of the first post-natal month he showed the typical Moro reflex (4) in response to a slap on the bed, at the end of the second month the reflex was still present though there was less bowing of the arms over the chest, and that at the end of the third month the gross reflex had diminished and the child gave only a general body jerk to this type of stimulus. In this case the data are longitudinal, recorded serially upon the same child, and the records embrace all the significant changes which take place in that particular characteristic. But intellectually there is very little one can do with such data other than repeat the descriptive statements of which the data are composed.

It has been repeatedly emphasized in this discussion that it is not so much the growing phenomenon *per se* but rather the mathematical demands imposed upon the data which determine the interval and limits of observations. At the same time it should be recognized that the acceptance of mathematical demands instead of the status of the existing phenomenon as the criterion for observational procedures is entirely justified because symbolic data, once they are organized into a meaningful system, can be manipulated on their own account so that relations and meanings can be educed from these manipulations which perhaps could not have been detected in the direct data. It is, indeed, in this sense that mathematics becomes the true handmaiden of scientific enquiry, not merely an elegant gown in which an investigation is dressed for public appearance.

One does not need to look to science in order to appreciate the advantage of an ordered symbolic system in dealing with an actual situation. Any motorist who struggled with the old descriptive Blue Books, less than a quarter of a century ago,

knows the value of modern road maps and the numbered highways. It is due not merely to the fact that highways have been expanded and improved but also to the ordering of highway information to a symbolic system that present day motoring can be carried on with facility. Again it should be emphasized that the reason for ordering observational or direct data in a symbolic system is because the symbols can be manipulated within their own system, independent of the actual phenomenon.

In emphasizing the influence of a symbolic order for determining the manner and frequency at which observations or measurements of a growing phenomenon should be made, the point should also be stressed that the original data may and often should lead to the formation of a new symbolic order. Mathematical systems are not finished to the extent that additional symbolic orders are precluded. Woodger (**6**) and Lewin (**3**) are both working on new symbolic systems for the ordering of biological data, and while the writer is not qualified to evaluate these methods, the fact that work is being done in that direction is reason for encouragement.

REFERENCES

1. Bertalanffy, L. von. *Modern theories of development.* Oxford University Press, 1933. Pp. 64–65.
2. Dewey, J. *Logic: the theory of inquiry.* New York: Henry Holt, 1938. P. 214.
3. Lewin, K. *Principles of topological psychology.* New York: McGraw-Hill Publ. in Psychology, 1936.
4. McGraw, Myrtle B. The Moro reflex. *Amer. J. Dis. Children,* 1937, **54**, pp. 240–251.
5. Palmer, C. E. *The analysis of longitudinal data.* Paper presented at the Eastern Regional Conference of the Society for Research in Child Development, New York City, December 4, 1937 (unpublished).
6. Woodger, J. H. *The axiomatic method in biology.* Cambridge University Press, 1937.

[MS. received September 14, 1939]

THE INADEQUACY OF THE KELLOGG–BRITT DEFINITION OF LEARNING

BY MATTHEW N. CHAPPELL

The Psychological Corporation

In the March 1939 issue of this publication another paper (7) in the Cason and Kellogg-Britt series (1–7) appeared in which Britt and Kellogg declare, "There is surely no such thing as 'one correct and perfect definition of learning' . . ." and argue again for the acceptance of their own definition. Learning is 'precisely' defined, they assert, as "a persisting change or modification of behavior which results from the activity of the organism itself, and which tends to adapt the organism to its environment."

In 1938, the present author pointed out to Kellogg in a private communication that his definition is inadequate for several obvious reasons. Kellogg replied that he was not in a position to discuss the matter in correspondence and suggested that I write an article to which he would then reply. Such a course seemed unnecessary at the time. However, since Kellogg and Britt in the recent article referred to the work of Chappell and Pike as supporting evidence for their position, it seems advisable now.

Let us first consider the separate parts of their definition, working backward from the end to the beginning. Since 'adapt' means 'make fit,' it would follow from the last phrase of the definition that modifications of behavior which are non-adaptive or counter to the organism's adaptation to its environment are not learned reactions. For example, when an individual dwells on gruesome or fearful thoughts for many hours a day over a period of weeks and subsequently manifests modifications of his behavior to such an extent that the thought of approaching a window on the fifteenth floor of a building will elicit in him feelings of panic, he could not be

conceived to have learned this activity. For by no stretch of the imagination could the modification of his emotional activity be considered to fit the man to his environment.

The middle phrase, "which results from the activity of the organism itself," implies that the activity itself is not the learning but rather that 'learning' is an abstract noun which refers in some way to the results of activity. Elsewhere in their article the authors indicate that they do regard learning as identical with the activities, since they argue for an operational conception which requires such identity. For example, in the operational definition of Force as the product of Mass and Acceleration,

$$F = MA,$$

Force is not a result of a function of Mass and Acceleration. It is identical with it. They are one and the same.

But, then again, this cannot be the Kellogg-Britt view of activity and learning, for the first clause of their definition holds that "learning is a persisting change or modification of behavior." There could be learning only so long as the operations which define it are operating according to the operational conception. Objects manifest force only so long as they are accelerating. So learning could apply only to the operations with which it is identical and would cease to be present in their absence.

Now consider the inadequacy of their whole definition in the light of our knowledge of learning.

At a given rate of practice, we know that some level of proficiency is eventually achieved at which further practice at that rate produces no further modification of behavior. Thus a pianist may practise an hour a day year after year and beyond a certain point manifest no increased skill. According to Kellogg and Britt, the daily hour of practice does not fall within the realm of learning after the close of the period of improvement. However, if the practice is stopped there is a measurable loss of skill in such circumstances.

Again, it is observed in many cases that modification of behavior occurs, if at all, only after a long series of influences

have operated on the individual or only after prolonged practice. Thus a rat may run a given maze once a week throughout his life and never modify his behavior so as to run the maze without an error. Does the trial that he takes each week not come within the scope of learning? If one trial does not, do two or a dozen?

Or, if as is frequently the case with us at the Psychological Corporation, we give a business executive evidence that a certain policy is antisocial and harmful to his business relations and he makes no change in his policy, can he be said to have learned nothing? We may subsequently give him evidence on four other occasions and still produce no modification. Later, as a result of a sixth bit of evidence, he may modify his behavior to the extent of changing the business policy. Did the first five experiences have nothing to do with the manifest change in the executive's behavior?

And learning plateaux, what of them? The typist on a 'plateau' between rises on her learning curve practices day after day with no modification of speed or accuracy. The practice she gives herself while she is on the plateau involves no learning, according to Kellogg's and Britt's definition, but the same practice during periods of measured gain does.

The most obvious objection to the definition is that it identifies learning with forgetting. Forgetting causes modifications of behavior that persist quite as well as those that Britt and Kellogg would call learned. On any view of the nature of forgetting, it is related to activity of the organism itself. One current view accounts for it in terms of 'interpolated' activity. The older view holds it is an activity of the organism which is spontaneous, once the activities necessary for learning have occurred. And, certainly, modifications which occur through forgetting are quite as adaptive as those that are learned.

Other objections such as its inability to account for 'over learning' could be urged against the Britt-Kellogg definition but perhaps those presented above are sufficient to indicate its inadequacy.

It is not the object of this paper to add a new definition

of learning. Like most other psychologists, perhaps, the author has a profound interest in the subject, the more so since it appears to him that there are only two possible mental processes—learning and forgetting—and that therefore all behavior, except the native responses, must be accounted for within their limits.

Regarded as a process, learning involves far more than either Cason or Kellogg and Britt contend. To initiate the process, work must be done on the organism or on one of its systems which produces some effect or change. The effect is cumulative as work continues to be done and may eventually attain such proportions as to increase the organism's scope or magnitude of reaction. When the work ceases to be done on the organism, its effect is dissipated in time and if modification has been manifested, eventually it diminishes progressively.

In the complete understanding of the learning and forgetting processes, we can make measurement on two parts—the activities through which work is done on the system and the concomitant modifications in behavior, when and if they occur. No direct measurement can be made on the effects produced on the system, but that makes them of no less importance to the understanding of the total learning process.

In their discussion of the neural or structural changes in learning, these authors are not in their best form. They are prevented from concluding that the afferent, efferent and central portions of the nervous system have little or nothing to do with learning only by what they term 'scientific conservatism.' The possibility of such a conclusion suggests itself to them because it is observed that under non-biological conditions many parts of the nervous system can be excised in different animals and some doubtful learning may still be manifested. The fact is that, by using small single lesions and the Britt-Kellogg form of argument, it can be demonstrated that few parts of the nervous system have any function at all, in learning or any other activity.

The fallacy of this kind of reasoning was pointed out by Wundt (**8**) when he wrote "Extirpation experiments . . . are not expected to reveal anything of moment concerning the

functions of the parts destroyed, but rather to illustrate the possible functions of those that are left intact."

Adolph Meyer, commenting on the same type of reasoning, remarked that if one leg of a tripod is kicked out of position the tripod will fall to the ground but that does not signify that the tripod was standing on one leg.

As Kellogg and Britt remark, our knowledge of the nervous system is limited, but it is certainly less so than they suppose.

REFERENCES

1. Britt, S. H. The learning-remembering process—(A reply to Professor Cason). Psychol. Rev., 1937, **44**, 462–469.
2. Cason, H. The concepts of learning and memory. Psychol. Rev., 1937, **44**, 54–61.
3. ——. Dr. Kellogg on the definition of learning. Psychol. Rev., 1938, **45**, 101–105.
4. ——. Dr. Britt on the concepts of learning, retention and reproduction. Psychol. Rev., 1938, **45**, 441–443.
5. Chappell, M. N., & Pike, F. H. Summation of stimuli and the neural changes in learning. Psychol. Rev., 1936, **43**, 283–307.
6. Kellogg, W. N. Some objections to Professor Cason's definition of learning. Psychol. Rev., 1938, **45**, 96–100.
7. Kellogg, W. N., & Britt, S. H. Structure or function in the definition of learning? Psychol. Rev., 1939, **46**, 186–197.
8. Wundt, W. *Principles of physiological psychology*. New York: Macmillan (Trans. by Tichener), 1910, p. 266.

[MS. received October 5, 1939]

THE SUPERFLUITY OF THE CHAPPELL CRITIQUE—A REPLY

BY W. N. KELLOGG

Indiana University

I am both complimented and surprised by the unusual amount of attention which Dr. Chappell has devoted to a few words written by Britt and myself, and I wish to thank him for the effort he has expended as a result of them. I have hesitated, nevertheless, about replying to his article and have given the matter a good deal of thought before finally deciding to do so. Sometimes I think that we psychologists make ourselves not a little ridiculous by our minute and attenuated arguments about this or that interpretation or result. To the disinterested observer outside the field of psychology, and to many psychologists themselves, controversies such as these are simply verbiage and usually end by convincing nobody. Despite such misgivings I have succumbed to temptation and am willing to be led once more into a debate which I fear by now has about worn itself out.

The thing which puzzles me most is that I cannot quite be sure what Dr. Chappell is arguing about. It looks to me as though he has missed the whole point of the Cason-Britt-Kellogg discussion. That point was clearly not to push the Kellogg or Britt definitions of learning, as Chappell's paper seems to assume. Rather it was to show that learning at the present time cannot be adequately defined in terms of the nervous system. We proposed our own definitions, in the spirit of fairness and openmindedness, because we thought that there ought to be constructive suggestions along with the critical or destructive ones which we offered. Certainly better definitions of learning are possible than any Britt or I have been able to formulate. My opinion was, and still is, however, that any one of our own definitions (and we suggested at

least three) was better than a strictly neurological definition of learning. The fundamental question here is what sort of a definition *best fits* the known facts. None can fit them all perfectly.

I would like it clearly understood, therefore, that I am not defending my previous statement because I think that it is the ultimate in definitions of learning; in fact, I do not think so. Neither have I previously tried, nor am I now trying, to 'sell' it to anybody. But the plain facts of the matter are that Chappell's point-by-point objections seem to me to be so very easily answered that I have not been able to refrain from doing so. Indeed I fail to see one really serious objection in the lot. It is largely for the sake of academic argument, therefore, rather than as a do-or-die defender of the faith, that I reply herewith to the Chappell objections.

Chappell's third paragraph. Even a person in this condition can be thought of as adapting to his environment if his experiences (thoughts, emotions, et cetera) are considered to be part of his 'internal environment.' His solution might then be regarded as adequate to his own internal and external environment.

Chappell's fourth paragraph. The modification or change in behavior is the learning.

Chappell's fifth paragraph. Learning is more like acceleration than force. It is the change which takes place.

Chappell's paragraphs 6 and 7. Correct. The pianist is not learning anything if his performance does not change. He is practicing but not learning.

Chappell's paragraphs 8 and 9. If a rat makes the same errors all the time he is not learning.

Chappell's paragraph 10. Of course. The man did not learn until he changed his behavior. No one knows whether the first 5 repetitions had any effect on the final response. In an illustration like the present one this is pure speculation.

Chappell's eleventh paragraph. Exactly. There would be no learning during the period of the plateau, providing the plateau was parallel to the baseline.

Chappell's twelfth paragraph. Perhaps forgetting is nega-

tive learning. At least this is a stimulating possibility. Perhaps also there is only one 'mental process,' namely learning, instead of the learning and forgetting which are postulated by Chappell.

In conclusion I would emphasize that there are two general ways of defining learning. These are the structural and functional ways. Britt and I happen to believe that there is not enough scientific data at the present time to justify a definition of learning which is based exclusively on structural nerve changes. A great many other psychologists believe the same thing. As far as I am concerned, that is all there is to it.